Über die Legitimation der Einführung künstlich intelligenter Systeme

Eine Analyse politischer Rechtfertigung und Kritik

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie (Dr. phil.)
durch die Philosophische Fakultät der
Heinrich-Heine-Universität Düsseldorf

Vorgelegt von:

Birte Keller

geboren in Düsseldorf

Erstgutachter: Prof. Dr. Frank Marcinkowski

Zweitgutachter: Prof. Dr. Christoph Bieber

Düsseldorf, Dezember 2025

D61

Bei dem vorliegenden Dokument handelt es sich um eine geringfügig angepasste Version der im Dezember 2025 unter gleichnamigem Titel eingereichten Inaugural-Dissertation zur Erlangung des Doktorgrades in Philosophie (Dr. phil.) an der Philosophischen Fakultät der Heinrich-Heine-Universität Düsseldorf.

Danksagung

An erster Stelle gilt mein Dank Prof. Dr. Frank Marcinkowski, an dessen Lehrstuhl ich seit bereits sieben Jahren einen Ort für meine akademische Entwicklung gefunden habe. Besonders schätze ich das große Maß an Freiheit und Vertrauen sowie die verlässliche Unterstützung. Ebenso danke ich Prof. Dr. Christoph Bieber für die geduldige Begleitung meiner Dissertation.

Ein weiterer Dank gilt meinen (ehemaligen) Kolleg:innen am Institut für Sozialwissenschaften. Die zahlreichen Gespräche und die kollegiale Unterstützung haben diese Zeit in besonderer Weise geprägt. Besonders danke ich Gerrit Philipps und Golo Flaßhoff, dass ihr euch im Rahmen verschiedener Kolloquien immer wieder Zeit genommen habt, meine Arbeit kritisch zu diskutieren und neue Impulse einzubringen. Ebenso danke ich den Hilfskräften des Lehrstuhls, die mich bei der empirischen Untersuchung dieser Arbeit unterstützt haben.

Ein außerordentlicher Dank gilt darüber hinaus Dr. Marco Lünich, der mich ebenso lange begleitet, wie meine wissenschaftliche Laufbahn andauert. Marco, ich danke dir nicht nur für die unzähligen Stunden, die du mit dem Korrekturlesen dieser Arbeit verbracht hast, sondern auch dafür, dass du jederzeit ein offenes Ohr und einen klugen Rat für mich hast – fachlich wie persönlich. In den letzten Jahren durfte ich dadurch sehr viel von dir lernen.

Ich danke meinen Freund:innen dafür, dass ihr mir immer wieder gezeigt habt, dass es neben dieser Arbeit noch so vieles andere auf der Welt gibt, und mir damit geholfen habt, dieses Projekt mit dem nötigen Abstand zu betrachten. Ihr habt mich immer ermutigt, motiviert und unterstützt. Das gilt in besonderer Weise für dich, Janine. Du verstehst mich ohne viele Worte und erkennst genau, welche Gedankenspirale gerade vielleicht einmal aufgebrochen werden muss. Ich danke dir für die Unterstützung im gesamten Prozess dieser Arbeit und weit darüber hinaus.

Von Herzen danke ich Mama und Papa, Lena und Nils. Euer Vertrauen in mich ist größer als mein eigenes. Ihr habt mir ermöglicht, diesen Weg überhaupt erst zu gehen, habt stets kleine und große Motivationsanreize für mich und seid bedingungslos immer für mich da. Dieser Rückhalt ist keineswegs selbstverständlich und macht mich unendlich dankbar.

Mein größter Dank gilt schließlich dir, Andi. Ich danke dir von ganzem Herzen für deine Geduld, dein Verständnis und deine unermüdliche Unterstützung. Ganz besonders in den letzten Wochen, aber auch schon seit so vielen Jahren, hältst du mir den Rücken frei und glaubst immer an mich. Danke, dass du alle Höhen und Tiefen mit mir teilst und mir ermöglicht hast, diese Arbeit fertigzustellen.

Velbert, Dezember 2025

Zusammenfassung

Die Dissertation untersucht, wie politische Akteur:innen die Einführung von KI-Systemen im Bildungswesen legitimieren, obwohl deren Folgen mit erheblichen Unsicherheiten verbunden sind. KI erscheint zunehmend als solutionistisches Instrument gesellschaftlicher Problemlösung, während Fragen nach Grundrechten, Verantwortung und Entscheidungsmacht häufig erst nach sichtbaren Fehlentwicklungen in den Fokus rücken. Vor diesem Hintergrund fragt die Arbeit, auf welcher moralischen Grundlage der KI-Einsatz gerechtfertigt und kritisiert wird, welche Rolle solutionistische Rechtfertigungen spielen und welche Konflikte daraus entstehen.

Die Arbeit entwickelt einen konzeptionellen Rahmen, der die Soziologie der Konventionen mit Politisierungsforschung und Techniksoziologie verbindet. Empirisch basiert sie auf einer qualitativen Inhaltsanalyse der bildungspolitischen KI-Debatte in Deutschland. Die Analyse zeigt eine technikoptimistische, aber diffuse Debatte. KI erscheint als unumgängliche Zukunftstechnologie, die Effizienz, Wettbewerbsfähigkeit und Fortschritt verspricht. Ansprüche an Gemeinwohl und Gerechtigkeit werden häufig in funktionale Kompromisse eingebunden. Dadurch wird primär die Ausgestaltung, kaum aber die Sinnhaftigkeit des KI-Einsatzes verhandelt. Kritik an Diskriminierung oder gefährdeter pädagogischer Sozialität bleibt randständig. Grundlegende Fragen nach Verantwortung, Entscheidungsmacht und gesellschaftlicher Ordnung im KI-Zeitalter werden somit weitgehend ausgeblendet.

Abstract

This dissertation examines how political actors legitimize the introduction of AI systems in education, despite substantial uncertainty about their consequences. AI increasingly appears as a solutionist instrument for addressing social problems, while questions of fundamental rights, responsibility, and decision-making power often come into focus only after visible failures have occurred. Against this background, the study asks on what moral grounds the use of AI is justified and criticized, what role solutionist justifications play, and what conflicts emerge.

The dissertation develops a conceptual framework that links the sociology of conventions with research on politicization and science and technology studies. Empirically, it is based on a qualitative content analysis of the German education-policy debate on AI. The analysis reveals a techno-optimistic yet diffuse debate. AI is framed as an inevitable technology of the future that promises efficiency, competitiveness, and progress. Claims to the common good and justice are frequently incorporated into functional compromises. As a result, the debate primarily addresses how AI should be implemented, while the question of whether it should be used at all is rarely negotiated. Criticism of discrimination or threats to pedagogical sociality remains marginal. Fundamental questions of responsibility, decision-making power, and social order in the age of AI are thus largely left unaddressed.

Inhaltsverzeichnis

1. Einführung **1**

2. Künstliche Intelligenz als soziotechnisches System verstehen **7**

2.1. Der Aufschwung Künstlicher Intelligenz im Bildungssektor als Ausdruck neoliberaler Restrukturierung 10

2.2. Potenziale Künstlicher Intelligenz für Gesellschaft und Bildung 13

2.3. Risiken Künstlicher Intelligenz für Gesellschaft und Bildung 15

2.4. Das Versprechen menschenzentrierter KI-Entwicklung zur Sicherung legitimer Künstlicher Intelligenz 18

2.5. Forderung einer Refokussierung der Forschung und politischen Steuerung 22

3. Theoretische Grundlagen politischer Legitimation im Kontext der Einführung Künstlicher Intelligenz **26**

3.1. Zur Abgrenzung statischer Legitimität und dynamischer Legitimation 27

3.2. Forschungsperspektiven auf die Mechanismen der Legitimitätsherstellung 28

3.3. Das pragmatische Legitimationsverständnis: Grundlagen der Soziologie der Konventionen 31

4. Politisierung als notwendige Bedingung legitimer Entscheidungen **34**

4.1. Politisierung als kontingenter politischer Aushandlungsprozess 35

4.2. Der graduelle demokratische Prozess zwischen Depolitisierung und (Re-)Politisierung 38

4.2.1. Funktionen und Dysfunktionen der Depolitisierung 39

4.2.2. Repolitisierung als Konfliktsymptom 40

4.2.3. Das Politisierungspotenzial Künstlicher Intelligenz 41

4.3. Drei Dimensionen der (De-)Politisierung 42

5. Zur Rolle der Rechtfertigungen in Legitimationsprozessen **44**

5.1. Cités als moralische Grundlage der Rechtfertigung und Kritik 46

5.1.1. Cité des Staatsbürgers 48

5.1.2. Cité des Marktes 49

5.1.3. Cité der Industrie 50

5.1.4. Cité der Inspiration 50

5.1.5. Cité des Hauses 51

5.1.6. Cité der Meinung 52

5.1.7. Cité des Projekts 52

5.1.8. Cité der Ökologie 53

5.1.9. Cité der Solution 54

5.2. Kompromisse als Ziel einer legitimen Entscheidung 55

5.2.1. Kritik als konstitutiver Bestandteil der Aushandlung legitimer Entscheidugnen 57

5.2.2. Drei Prüfungssituationen der Aushandlung legitimer Entscheidungen 58

5.3. Konzeptioneller Rahmen der Analyse der Legitimation der politischen Entscheidung über den Einsatz künstlicher Intelligenz 62

6. Techniksoziologische Deutungsmuster als Legitimationsgrundlage moderner Technologien 64

6.1. Das unausweichliche Fortschrittsversprechen der Technik: Die Perspektive des optimistischen Determinismus 68

6.1.1. Rechtfertigung technologischer Innovation zwischen Effizienz, Wettbewerb und Allheilmittel 69

6.1.1.1. Der Aufschwung des Solutionismus: Künstliche Intelligenz gerechtfertigt als Allheilmittel 70

6.1.1.2. Die Rechtfertigung Künstlicher Intelligenz im Kontext des internationalen Wettbewerbs und Wohlstandsversprechen 72

6.1.2. AI as „political-administrative necessity“ – Mangelnde Politisierung aufgrund der Unausweichlichkeit Künstlicher Intelligenz 73

6.2. Der Kontrollverlust der Technik: Die Perspektive des pessimistischen Determinismus 74

6.2.1. Technikkritik: Das Legitimationsdefizit durch technischen Kontrollverlust und Technokratisierung 77

6.2.1.1. Künstliche Intelligenz als Instrument der Kontrolle und Überwachung 78

6.2.1.2. Das technokratische Repräsentationsdefizit künstlich intelligenter Entscheidungen 79

6.2.2. Der rekursive Politisierungsprozess: Technikkritik als Ausdruck reaktiver Politisierung und als Folge technischer Fehlschläge 80

6.3. Die Demokratisierung technischer Gestaltung als Folge der Anerkennung der Ambivalenz von Technik: Die Perspektive des Sozialkonstruktivismus 82

6.3.1. Die Anerkennung der politischen Qualität und Ambivalenz von Technik 82

6.3.2. Die techniksoziologische Kritik der Verschleierung der sozialen Konstruktion von Technik 84

6.3.3. Partizipation als politische Rechtfertigung einer demokratischen Technikentwicklung 86

6.3.3.1. Die Chance der aktiven demokratischen Gestaltung des Einsatzes Künstlicher Intelligenz 86

6.3.3.2. Investitionen in Bildung und Forschung zur Sicherung wirtschaftlicher und gesellschaftlicher Teilhabe 87

6.3.4. Fallstricke einer beabsichtigten Politisierung durch partizipative Technikgestaltung 88

6.4. Die moralische Fundierung der Technikgestaltung: Die Perspektive des kritischen Pragmatismus 89

6.4.1. Das pragmatische Verständnis der moralischen Konstruktion Künstlicher Intelligenz 91

6.4.2. Technik als Träger von Moral und Machtverhältnissen und die Notwendigkeit der Kritik 93
6.4.3. Die politische Rechtfertigung einer wertebasierten Entwicklung Künstlicher Intelligenz 94
6.4.4. Zwischen Wertorientierung und technokratischer Verkürzung gesellschaftlicher Herausforderungen: Die techniksoziologische Kritik des Solutionismus 95
6.4.4.1. Die Kritik an der Verkürzung der moralischen Gestaltung Künstlicher Intelligenz 95
6.4.4.2. Unsachgemäße Komplexitätsreduktion: Die Kritik der solutionistischen Rechtfertigung 97
6.4.5. Die Forderung der Politisierung Künstlicher Intelligenz aus technopragmatischer Perspektive 99

7. Lehren und Desiderate der techniksoziologischen und technologiepolitischen Auseinandersetzung für die Legitimation Künstlicher Intelligenz 101

7.1. Der Bedarf der Rekonstruktion der moralischen Grundlagen politischer Rechtfertigungen Künstlicher Intelligenz 103
7.2. Die Analyse der Rolle solutionistischer Rechtfertigungen für die Legitimation Künstlicher Intelligenz 106
7.3. Die Untersuchung der kritischen Fundierung der Legitimation Künstlicher Intelligenz 108

8. Methodisches Vorgehen 110

8.1. Definition des Untersuchungsgegenstands 111
8.2. Dokumentenrecherche und -auswahl 112
8.2.1. Strukturierung der Forschungsfrage 114
8.2.2. Definition von Einschlusskriterien: Untersuchungszeitraum, Datenbanken und Suchbegriffe 114
8.2.3. Dokumentenrecherche 116
8.2.4. Dokumentenauswahl 117
8.3. Inhaltlich Strukturierende Qualitative Inhaltsanalyse 120
8.3.1. Vorgehen 121
8.3.2. Kategoriensystem 124

9. Ergebnisse 126

9.1. Kontexte der bildungspolitischen Debatte über Künstliche Intelligenz 127
9.1.1. Differenzierung der thematisierten Anwendungen Künstlicher Intelligenz 128
9.1.2. Diffuse Vorstellungen des Einsatzes Künstlicher Intelligenz im Bildungswesen 130
9.1.3. Wissenserwerb über Künstliche Intelligenz 131
9.1.4. Forschung über Künstliche Intelligenz 132
9.1.5. Investitionen in Künstliche Intelligenz 133

9.2. Akteur:innen der bildungspolitischen Debatte über Künstliche Intelligenz 134
9.3. Die moralische Fundierung der Rechtfertigung des Einsatzes Künstlicher Intelligenz 136
9.3.1. Das industrielle Streben nach Effizienz, Planbarkeit und Innovation 140
9.3.1.1. Effizientere und optimierte Bildung durch den Einsatz Künstlicher Intelligenz 140
9.3.1.2. Planungssicherheit und Evaluation durch die Standardisierung künstlicher Intelligenz 142
9.3.1.3. Künstliche Intelligenz als Treiber eines innovativen Bildungssystems und Forderungen nach Investitionen und Fachkräfteentwicklung 142
9.3.2. Die Zwänge des Wettbewerbs und die Chancen der Individualisierung aus Sicht der Cité des Marktes 143
9.3.2.1. Künstliche Intelligenz als Garant der Sicherung der Wettbewerbsfähigkeit von Bund, Ländern und Hochschulen 144
9.3.2.2. Internationale Vernetzung und Fachkräfteausbildung als Voraussetzung der Wohlstandsgewinne durch den Einsatz Künstlicher Intelligenz 146
9.3.2.3. Individualisierung der Bildung als persönlicher Wettbewerbsvorteil 147
9.3.3. Staatsbürgerliche Gemeinwohlorientierung als legitimatorische Grundlage des Einsatzes Künstlicher Intelligenz 148
9.3.3.1. Künstliche Intelligenz als Instrument der Gemeinwohlförderung auf Basis ethischer Prinzipien 148
9.3.3.2. Teilhabe und Mitbestimmung als staatsbürgerliche Rechtfertigung eines legitimen Einsatzes Künstlicher Intelligenz 151
9.3.3.3. Inklusion und Chancengleichheit als staatsbürgerliche Rechtfertigung des Einsatzes Künstlicher Intelligenz und als Kompromissgrundlage industrieller und marktorientierter Ansprüche 154
9.3.4. Die Vernetzung des Bildungssektors aus Sicht der Cité des Projekts 156
9.3.4.1. Künstliche Intelligenz als Projekt moderner Gesellschaften 156
9.3.4.2. Netzwerkbildung als Flexibilisierungsanforderung 157
9.3.5. Künstliche Intelligenz als Imagegewinn und das Akzeptanzproblem der Cité der Meinung 159
9.3.5.1. Künstliche Intelligenz als Instrument des Reputationsgewinns 159
9.3.5.2. Akzeptanz als notwendige Voraussetzung des Einsatzes Künstlicher Intelligenz 160
9.4. Zur Rolle solutionistischer Rechtfertigungen des Einsatzes Künstlicher Intelligenz 162
9.4.1. Die Unvermeidbarkeit Künstlicher Intelligenz als Ausdruck des optimistischen Determinismus solutionistischer Rechtfertigungen 163

9.4.2. Künstliche Intelligenz als Problemlösungsinstrument und Verbesserung des Status quo 164
9.4.3. Die Diffusität solutionistischer Rechtfertigungen als Ausdruck eines digitalen Geists 166
9.4.4. Der innere Widerstreit des Solutionismus: eine Realitätsprüfung der Problemlösungsfähigkeit und Wirkkraft Künstlicher Intelligenz 167
9.4.5. Mangelnde Rechtfertigung als Ausdruck eines deterministischen Optimismus: Der implizite Solutionismus 170
9.4.5.1. Mangelnde Rechtfertigung der Investitionen in den Aufbau Künstlicher Intelligenz 170
9.4.5.2. Mangelnde Rechtfertigung des Kompetenzerwerbs 171
9.4.6. Effiziente Problemlöser: Der technikoptimistische Kompromiss von Industrie und Solution 173
9.4.7. Wertebasierte Problemlösung: Ein instrumenteller Kompromiss von Staatsbürger und Solution 175
9.5. Kritik und Konfliktlinien der moralischen Rechtfertigung des Einsatzes Künstlicher Intelligenz 178
9.5.1. Staatsbürgerliche Kritik der Verletzung von Grundrechten und demokratischer Werte durch Künstliche Intelligenz 178
9.5.1.1. Die Gefährdung des Gerechtigkeitsanspruchs durch mangelnde Objektivität und drohende Diskriminierung 179
9.5.1.2. Die Realitätsprüfung des humanistischen Bildungsverständnisses im Kontext des Einsatzes Künstlicher Intelligenz 182
9.5.1.3. Die Verletzung von Persönlichkeitsrechten durch den Einsatz Künstlicher Intelligenz 189
9.5.1.4. Unsicherheit über Technikfolgen und Regulierungsbedarfe als Prüfstein staatsbürgerlicher Legitimation 197
9.5.2. Häusliche Kritik der Gefährdung des sozialen Bildungsprozesses durch Künstliche Intelligenz: Vertrauen auf das Tradierte 200
9.5.2.1. Die häusliche Kritik an der Abwertung des pädagogischen Selbstverständnisses im Konflikt mit solutionistischer Steuerung und industrieller Standardisierung 201
9.5.2.2. Die häusliche Kritik des Arbeitsplatzverlustes im Konflikt mit industrieller Rationalisierung und solutionistischem Fortschrittsdenken 202
9.5.2.3. Die Aushöhlung der Sozialität des Bildungsprozesses durch solutionistische und industrielle Deutungsmuster 204
9.5.2.4. Die Diskrepanz zwischen Pädagogik und Datenerfassung im Konflikt mit industrieller Output-Orientierung 207
9.5.2.5. Die häusliche Kritik des Verlusts tradierter Kompetenzen im Konflikt mit industrieller Rationalisierung und solutionistischer KI-Ausbildung 209

9.6. Zusammenfassung der Analyseergebnisse 210
9.6.1. Die moralische Grundlage der politischen Rechtfertigung des Einsatzes Künstlicher Intelligenz 211
9.6.2. Zur Rolle solutionistischer Rechtfertigungen in der politischen Rechtfertigung des Einsatzes künstlicher Intelligenz 213
9.6.3. Die moralische Grundlage der politischen Kritik des Einsatzes Künstlicher Intelligenz 214
10. Diskussion 216
10.1. Konsequenzen der Diffusität der politischen Debatte über Künstliche Intelligenz 216
10.1.1. Symbolpolitik als Konsequenz eines Hypes um Künstliche Intelligenz und als Mechanismus kommunikativer Stabilisierung politischer Handlungsfähigkeit 217
10.1.2. Mangelnde politische Verantwortung aufgrund des defizitären Verständnisses Künstlicher Intelligenz 220
10.2. Politisierung Künstlicher Intelligenz im Spannungsfeld von Kontingenz und Konflikt 221
10.2.1. Die Einschränkung der Kontingenz als Revival des techniksoziologischen Dualismus? 222
10.2.2. Konfliktdynamiken und Prüfungsformen: Zwischen Utopie, Dystopie, Realitätsprüfungen und radikaler Kritik der Cité des Hauses 224
10.2.3. Die Depolitisierung des Solutionismus: Performative Konsequenzen eines optimistischen Determinismus 227
10.2.4. Die Cité des Staatsbürgers als heiliger Gral einer integrativen Politisierung Künstlicher Intelligenz 230
10.3. Konsequenzen für die Legitimation Künstlicher Intelligenz 232
10.3.1. Politisierung als notwendige, nicht jedoch als hinreichende Bedingung legitimer Entscheidungsfindung 233
10.3.2. Die fragile Legitimität symbolischer Politik 234
10.3.3. Verpasstes Momentum? Politische Legitimation und die Zukunft der Gestaltung Künstlicher Intelligenz 237
11. Fazit 238
11.1. Lehren der Arbeit für Wissenschaft, Politik und Zivilgesellschaft 241
11.1.1. Reflexion der wissenschaftlichen Erkenntnisse 241
11.1.2. Reflexion der Erkenntnisse für die politische Praxis 242
11.1.3. Reflexion der Erkenntnisse aus Sicht der Zivilgesellschaft 244
11.2. Limitationen und anschließende Forschungsbedarfe 245
11.3. Ausblick 247
Literaturverzeichnis 250
Anhang 297
Eidesstattliche Versicherung 346

Abkürzungsverzeichnis

BT	Bundestag
BR	Bundesrat
DFKI	Deutsches Forschungszentrum für Künstliche Intelligenz
DIP	Dokumentations- und Informationssystem für Parlamentsmaterialien
DL	Deutscher Lehrerverband
DSGVO	Datenschutzgrundverordnung
EFI	Expertenkommission Forschung und Innovation
EU	Europäische Union
FF	Forschungsfrage
GEW	Gewerkschaft Erziehung und Wissenschaft
GWK	Gemeinsame Wissenschaftskonferenz
Herv. i. O.	Hervorhebung im Original
HFD	Hochschulforum Digitalisierung
HLB	Hochschullehrerbund
HRK	Hochschulrektorenkonferenz
KI	Künstliche Intelligenz
KMK	Kultusministerkonferenz
LT	Landtag
ML	Machine Learning/Maschinelles Lernen
SdK	Soziologie der Konventionen
STS	Science and Technology Studies
TAB	Büro für Technikfolgenabschätzung
WR	Wissenschaftsrat

Abbildungsverzeichnis

Abbildung 1. Konzeptioneller Rahmen zur Analyse legitimer politischer Entscheidungsfindung des Einsatzes künstlicher Intelligenz 64

Abbildung 2. Flowchart zur Auswahl der Untersuchungseinheiten in Anlehnung an Moher et al. (2009) 113

Abbildung 3. Vorgehen der inhaltlich strukturierenden qualitativen Inhaltsanalyse 122

Abbildung 4. Relative Häufigkeitsverteilung der Cités auf Segmentebene 137

Abbildung 5. Valenz der Rechtfertigungen nach Cités 138

Abbildung 6. Relative Häufigkeit der Prüfungsformen nach Cités 140

Tabellenverzeichnis

Tabelle 1. Übersicht der drei Prüfungsformen der Aushandlung nach Boltanski (2011) 59

Tabelle 2. Techniksoziologische Deutungsmuster, ihre Kritik- und Rechtfertigungsmuster und ihre Konsequenzen für die Politisierung von Technik 67

Tabelle 3: Anwendung der PICOC-Methode nach Petticrew & Roberts (2006) 114

Tabelle 4: Suchbegriffe der Dokumentenrecherche 116

Tabelle 5. Suchabfragen, Suchtreffer und Untersuchungseinheiten nach Institutionen 118

Tabelle 6. Übersicht thematisierter KI-Systeme nach Anwendungsbereichen 129

1. Einführung

Im Jahr 2015 prophezeiten Russell und Kollegen: „Success in the quest for artificial intelligence has the potential to bring unprecedented benefits to humanity“ (S. 105). Heute, zehn Jahre später, scheint dieses Versprechen in der Breite der Gesellschaft angekommen zu sein. Anwendungen künstlicher Intelligenz (KI) werden gegenwärtig in nahezu allen Lebensbereichen als „magische“ Problemlöser der Neuzeit präsentiert (Elish & boyd, 2018). Das Potenzial der KI scheint grenzenlos zu sein und reicht von der Reduktion monotoner Arbeitsprozesse bei gleichzeitiger Steigerung von Effizienz, Effektivität und Präzision bis hin zu einer vermeintlich höheren Neutralität KI-basierter Entscheidungsfindung (Engin & Treleaven, 2019; Naseem et al., 2020; Xu, 2022). KI-Anwendungen tragen zum Beispiel zur Früherkennung von Krankheiten bei (Alowais et al., 2023; Deo, 2015), versprechen eine objektive Bewertung in der Rechtsprechung (Xu, 2022) oder der Kreditwürdigkeit von Privatpersonen (Cao, 2023; Lessmann et al., 2015), ermöglichen die Individualisierung von Lernprozessen (Chiu et al., 2023; Zawacki-Richter et al., 2019) oder unterstützen die Bearbeitung moderner Krisen wie Pandemien oder Klimawandel, etwa durch die Prognose von Infektionsverläufen oder Extremwetterereignissen (Cowls et al., 2023; Malik et al., 2021; Naseem et al., 2020; Vaishya et al., 2020).

Während sich dieses optimistische Bild eines digitalen Problemlösers nicht nur in einem flächendeckenden KI-Einsatz widerspiegelt, sondern ebenso in einer überwiegend positiven medialen KI-Berichterstattung (Brennen et al., 2018; Chuan et al., 2019; Fast & Horvitz, 2017; Fischer & Puschmann, 2021; Köstler & Ossewaarde, 2022; Vergeer, 2020; Zeng et al., 2022), wird die Entwicklung, KI in allen erdenklichen Lebens- und Gesellschaftsbereichen einsetzen zu wollen, in der wissenschaftlichen Auseinandersetzung insbesondere aus techniksoziologischer Perspektive zunehmend als solutionistisch kritisiert. Gemeint ist die fälschliche Annahme, nahezu jedes noch so komplexe gesellschaftliche Problem ließe sich durch den Einsatz von KI lösen (Bartl, 2023; Morozov, 2013; Obendiek & Seidl, 2023). Eine solche Sichtweise vernachlässigt jedoch sowohl bestehende Risiken als auch die Unsicherheit über unintendierte Folgen des KI-Einsatzes (Elish & boyd, 2018; Nordström, 2022). Zu diesen Risiken zählen unter anderem diskriminierende Entscheidungen, mangelnde Transparenz der zugrunde liegenden Verfahren, unzureichender Datenschutz und der Verlust menschlicher Autonomie in sensiblen Entscheidungsprozessen (Bartneck et al., 2021; Wirtz et al., 2022).

Wenngleich diese Risiken nur langsam in das öffentliche Bewusstsein vordringen (Brantner & Saurwein, 2021; Kieslich et al., 2023; Nguyen & Hekman, 2024; Schreiber & Schreiber, 2024), haben sie jedoch in jüngerer Zeit dazu beigetragen, dass im Rahmen der Technikregulierung

verstärkte Anstrengungen unternommen wurden, einen verantwortungsvollen Umgang mit KI zu gewährleisten. Neben dem bislang weltweit einzigartigen *AI Act* der Europäischen Union (EU, 2024) wurden zahlreiche ethische Leitlinien politischer, zivilgesellschaftlicher und wirtschaftlicher Akteur:innen veröffentlicht. Ziel ist es, eine menschenzentrierte und vertrauenswürdige Entwicklung von KI zu fördern, indem sie Prinzipien wie Fairness, Transparenz und Verantwortlichkeit betonen und die Bedeutung der Evaluation hervorheben, die eine fortlaufende Optimierung der Systeme ermöglichen soll (Corrêa et al., 2023; Hagendorff, 2020; Jobin et al., 2019; Shneiderman, 2020; Zweig et al., 2018). In der Praxis werden jedoch viele Fehlerquellen erst nach der Implementierung von KI-Systemen sichtbar, sodass bereits eingetretene Folgen mitunter gravierende Konsequenzen für Individuen und Gesellschaft nach sich ziehen.

Die Umsetzung der ethischen Richtlinien stößt damit auf das sogenannte *Collingridge-Dilemma* (Collingridge, 1985): In frühen Phasen der KI-Entwicklung erscheint eine angemessene Regulierung kaum möglich, da die konkreten Folgen der Technologie noch unklar sind und ihr Einfluss auf Individuen und Gesellschaft nicht verlässlich abgeschätzt werden kann. Sind KI-Anwendungen jedoch erst einmal flächendeckend implementiert, kommen regulatorische Eingriffe häufig zu spät und Korrekturen sind nur schwer umsetzbar, da die damit verbundenen Konsequenzen bereits manifest geworden sind (Awad et al., 2020; Kudina & Verbeek, 2019). Dieses Dilemma lässt sich in der Geschichte technologischer Entwicklungen von der Atomenergie bis zu sozialen Netzwerken beobachten. Im Fall der Atomkraft wurden Risiken wie Reaktorunfälle, langfristige Umwelt- und Gesundheitsschäden sowie ungelöste Fragen der Endlagerung erst nach dem Aufbau umfangreicher Infrastrukturen sichtbar. Die daraus resultierenden gesellschaftlichen Folgen werden in erheblichem Maße von der Bevölkerung getragen, etwa durch nukleare Katastrophen oder anhaltende soziale Konflikte im Zusammenhang mit der Endlagersuche. Politische Korrekturen sind zudem mit hohen finanziellen Belastungen verbunden, da Stilllegung, Rückbau und Entsorgung der Anlagen über Jahrzehnte hinweg erhebliche Kosten verursachen (von Hirschhausen et al., 2015). Auch bei sozialen Netzwerken traten Risiken wie Desinformation, psychische Belastungen, Datenschutzverletzungen und die Machtkonzentration großer Technologieunternehmen erst nach ihrer massenhaften Nutzung zutage. Aus Sorge um den Schutz junger Menschen vor psychischen Schäden und algorithmischer Beeinflussung wurde etwa in Australien kürzlich ein Verbot sozialer Medien für unter 16-Jährige eingeführt (Fatt & Fardouly, 2025). Regulierungsversuche stoßen dabei mitunter auf den Widerstand der Plattformunternehmen und gehen mit hohen Kosten für staatliche Aufsicht, gesellschaftliche Schadensbegrenzung und den Schutz demokratischer Prozesse einher (Bouza García & Oleart, 2024; Nocun, 2018; Stockmann, 2023). Vor diesem Hintergrund stellt sich

jenseits des gegenwärtigen KI-Hypes die Frage, weshalb KI-Anwendungen zunehmend flächendeckend in gesellschaftlichen Bereichen eingesetzt werden, obwohl historische Erfahrungen als Warnsignale verstanden werden können.

Wie gravierend sich diese Problematik bereits heute im Fall von KI auswirken kann, zeigen zahlreiche, teils diskriminierende Anwendungen in unterschiedlichen Handlungsfeldern. Neben Systemen der Strafverfolgung (Angwin et al., 2016), des Arbeitsmarktes (Allhutter et al., 2020; Dastin, 2022) oder des Finanzwesens (Telford, 2019) haben auch im Bildungswesen KI-basierte Anwendungen unbeabsichtigte und folgenreiche Effekte hervorgebracht. Ein prominentes Beispiel bot der Einsatz eines vom staatlichen *Office of Qualifications and Examinations Regulation* in Großbritannien entwickelten Algorithmus zur vermeintlich fairen Vergabe von Abiturnoten während der COVID-19-Pandemie. Dieser benachteiligte systematisch Schüler:innen aus sozioökonomisch schwächeren Familien gegenüber angehenden Absolvent:innen von Privatschulen (Edwards, 2021; Hao, 2020). Angesichts der Relevanz der Noten für Ausbildung und berufliche Laufbahn sowie massiver öffentlicher Proteste sah sich die Behörde gezwungen, das System zurückzuziehen (Smith, 2020).

Der Fall verdeutlicht exemplarisch, dass KI-Anwendungen häufig erst dann in den Fokus öffentlicher Aufmerksamkeit und Kritik geraten, wenn negative Folgen bereits eingetreten sind. Erst das Sichtbarwerden von Fehlentscheidungen und Diskriminierung führt zur Infragestellung des Systems (Mittelstadt et al., 2016). In diesem Zuge geraten nicht nur die technische Ausgestaltung, sondern auch die dem KI-Einsatz zugrunde liegenden sozialen Normen, wie im vorliegenden Beispiel der Schutz sozialer Gerechtigkeit, nachträglich in die Kritik (König & Wenzelburger, 2021; Ulbricht, 2020; Wieringa, 2023). Dieser Prozess lässt sich als Ex-post-Politisierung der Technologie verstehen, die auf ein Legitimationsdefizit ihrer Anwendung verweist. Durch die Infragestellung der KI-Anwendung und der ihr zugrunde liegenden sozialen Normen verliert der KI-Einsatz seinen selbstverständlichen Charakter und löst einen erhöhten Rechtfertigungsdruck auf politische Entscheidungsträger:innen und Anwender:innen von KI-Systemen aus. Sie müssen darlegen, auf welchen moralischen Maßstäben ihre Entscheidungen beruhen und warum der Einsatz von KI als legitim angesehen werden kann (Diaz-Bone & Schrör, 2022). Die Verletzung sozialer Normen wirkt sich dabei nicht nur auf die Bewertung der Technologie aus, sondern belastet zugleich das Vertrauen in die politischen Institutionen, die ihre Einführung verantworten (Schiff et al., 2022).

Vor diesem Hintergrund wird die Problematik des im Collingridge-Dilemma beschriebenen und in der Praxis häufig verfolgten Vorgehens des Trial-and-Error besonders deutlich. Bei

diesem Vorgehen werden KI-Systeme vielfach breit eingesetzt und erst im laufenden Betrieb korrigiert oder zurückgenommen, wenn bereits Schäden entstanden sind. Unbeabsichtigte Folgen und schwerwiegende Konsequenzen für Individuen oder ganze gesellschaftliche Bereiche, die erst nachträglich sichtbar werden, werden damit wissentlich in Kauf genommen. Dies steht etwa im Kontrast zu anderen Hochrisikotechnologien wie Arzneimitteln oder Impfstoffen, die vor ihrem Einsatz langjährige Prüf- und Zulassungsverfahren durchlaufen.

Die mit einem voreiligen KI-Einsatz verbundene Risikodynamik wirkt dabei auf mehreren Ebenen und verleiht der Frage eines demokratisch-legitimen Einsatzes von KI sowohl gesellschaftspolitisch als auch demokratietheoretisch besondere Relevanz. Gesellschaftspolitisch birgt die Implementierung algorithmischer Entscheidungssysteme erhebliche Risiken, insbesondere im Hinblick auf potenziell diskriminierende Auswirkungen für die von den Entscheidungen Betroffenen (Fazelpour & Danks, 2021; Giovanola & Tiribelli, 2022; Keller et al., 2022). Das Beispiel des britischen Benotungsalgorithmus verdeutlicht, wie solche Systeme bestehende soziale Ungleichheiten reproduzieren oder verstärken und dabei zentrale demokratische Prinzipien wie Chancengleichheit und Verfahrensgerechtigkeit verletzen können. Diese Effekte verweisen über den konkreten Anwendungsfall hinaus auf strukturelle Risiken des KI-Einsatzes in gesellschaftlich sensiblen Kontexten, insbesondere dort, wo automatisierte Entscheidungen individuelle Lebensverläufe beeinflussen. Die Risiken äußern sich unter anderem in mangelnder Akzeptanz algorithmischer Systeme (Shin, 2021; Wenzelburger et al., 2022) und in öffentlichem Protest (Edwards, 2021; Marcinkowski et al., 2020). Diese Reaktionen können über die Ablehnung spezifischer Anwendungen hinaus als Ausdruck eines grundlegenden Unbehagens gegenüber der zunehmenden Automatisierung gesellschaftlicher Entscheidungsprozesse verstanden werden. Die Kritik richtet sich dabei vor allem gegen die Verschiebung von Verantwortung, Entscheidungsmacht und normativer Bewertung von menschlichen Akteur:innen hin zu algorithmischen Verfahren (Crawford, 2021; Kalluri, 2020).

Aus demokratietheoretischer Perspektive verweist die Ablehnung algorithmischer Entscheidungssysteme auf ein Legitimationsdefizit, das insbesondere dann sichtbar wird, wenn Entscheidungen als intransparent, nicht anfechtbar oder sozial ungerecht wahrgenommen werden (Yeung, 2018). Der daraus resultierende Vertrauensverlust in Anwender:innen sowie in politische Entscheidungsträger:innen, die den Einsatz dieser Technologien ermöglichen oder vorantreiben (Chai et al., 2022; Schiff et al., 2022; Vertesi & boyd, 2023), kann den KI-Einsatz zum Katalysator einer breiteren Vertrauenskrise machen, in der technologische Innovation bestehende Spannungen zwischen Staat, Institutionen und Bürger:innen verschärft. Der Einsatz von KI in öffentlichen Entscheidungsprozessen stellt daher weniger eine technische als eine

politisch-gesellschaftliche Herausforderung dar. Er wirft grundlegende Fragen nach demokratischer Kontrolle und normativer Legitimation auf. Zwar wird KI von politischen Akteur:innen häufig als Instrument der Effizienzsteigerung oder als gesellschaftliches Wohlstandsversprechen dargestellt (Bareis & Katzenbach, 2022; Guenduez & Mettler, 2023; Hirsch-Kreinsen & Krokowski, 2024b; Köstler & Ossewaarde, 2022; Ulnicane et al., 2022), doch bleibt vielfach ungeklärt, auf welcher normativen Grundlage diese Systeme gesellschaftlich akzeptiert werden sollen und welche Werte sie priorisieren. Die Annahme, technologische Rationalität könne politische oder moralische Aushandlungsprozesse ersetzen, erweist sich dabei als problematisch. Langfristig droht der routinierte Einsatz algorithmischer Systeme, technokratische Entscheidungslogiken zu normalisieren, in denen demokratische Aushandlung, individuelle Ermessensspielräume und soziale Kontextualisierung an Bedeutung verlieren (Bartl, 2023; Sætra, 2020; Smithers, 2020). Die Folgen des KI-Einsatzes betreffen damit grundlegende Macht-, Verantwortungs- und Vertrauensverhältnisse in modernen Demokratien und erfordern eine umfassende gesellschaftliche und politische Auseinandersetzung über die Rolle automatisierter Systeme in kollektiven Entscheidungsprozessen.

Diese doppelte gesellschaftspolitische und demokratietheoretische Risikodynamik verweist auf die Problematik einer einseitigen Fokussierung der öffentlichen und politischen Debatte auf die Frage, *wie* KI-Systeme bestmöglich und flächendeckend eingesetzt werden können, ohne in den Blick zu nehmen, *ob* ihr Einsatz angesichts der bestehenden Risiken und der Unsicherheit über ihre Konsequenzen gerechtfertigt ist (Awad et al., 2018; Crawford, 2021; Smuha, 2022; Züger & Asghari, 2023). Während eine Ex-post-Politisierung nach Schadenseintritt die dem KI-Einsatz zugrunde liegenden moralischen Grundlagen offenlegt und zur Disposition stellt, wird zugleich deutlich, dass die Legitimation von KI-Systemen nicht erst nach ihrer Implementierung erfolgen kann. Angesichts der potenziell schwerwiegenden Folgen ist eine vorgelagerte politische Auseinandersetzung notwendig, in der die Grundlagen einer legitimen Entscheidung über den KI-Einsatz ausgehandelt werden. Eine Ex-ante-Politisierung kann daher als zentraler Bestandteil demokratischer Entscheidungsprozesse verstanden werden (Hay, 2014; Zürn, 2013). Sie eröffnet einen Raum, in dem Kritik und Rechtfertigung artikuliert werden können und in dem unterschiedliche moralische Grundsätze zur Diskussion stehen. Auf dieser Grundlage können politische Akteur:innen tragfähige Kompromisse entwickeln und eine demokratisch legitime Entscheidung für oder gegen den KI-Einsatz treffen (Boltanski & Thévenot, 1991/2018; Diaz-Bone, 2022).

Vor dem Hintergrund des zunehmend inflationären KI-Einsatzes in sensiblen gesellschaftlichen Bereichen und der Unsicherheit über die damit verbundenen potenziell schwerwiegenden

Konsequenzen verfolgt die vorliegende Arbeit das Ziel der empirischen Klärung, wie politische Entscheidungsträger:innen die Einführung künstlich intelligenter Systeme legitimieren, obwohl diese Technologie erhebliche Risiken birgt.

Die Beantwortung dieser Frage verspricht nicht nur vertiefte Erkenntnisse über die Konstruktion der Legitimität politischer Entscheidungen im Allgemeinen, sondern auch über die spezifische Rolle politischer und normativer Kritik- und Rechtfertigungsmuster in Legitimationsprozessen moderner Technologien. Eine Analyse der moralischen Grundlagen dieser Muster erlaubt zudem, Einsichten in gegenwärtige gesellschaftliche und politische Herausforderungen sowie in potenzielle Konfliktlinien des digitalen Zeitalters zu gewinnen. Dabei liegt das Augenmerk auf dem Bildungssektor, der in dieser Arbeit als exemplarischer Anwendungsfall der Analyse dient. In diesem Kontext können KI-basierte Entscheidungen nicht nur direkte Konsequenzen für die individuelle Entwicklung insbesondere junger Menschen und ihre Bildungschancen haben, sondern auch weitreichende gesellschaftliche Auswirkungen nach sich ziehen, etwa im Hinblick auf soziale Ungleichheit und Teilhabe. Zugleich sind Bildungsinstitutionen als öffentliche Einrichtungen unmittelbar durch politische Entscheidungen über den KI-Einsatz betroffen. Für die wissenschaftliche Debatte leistet die vorliegende Arbeit damit einen Beitrag zur Weiterentwicklung des Verständnisses von Legitimität in technologiepolitischen Kontexten. Für die politische Praxis bietet sie Orientierungswissen, wie zukünftige Technologieeinführungen so gestaltet werden können, dass sie nicht nur effizient, sondern auch demokratisch tragfähig sind. Für die Zivilgesellschaft eröffnet sie schließlich die Möglichkeit, Mechanismen politischer Rechtfertigung und Kritik besser zu verstehen, Macht- und Entscheidungsprozesse transparenter nachzuvollziehen und dadurch die eigene Teilhabe an technologiepolitischen Aushandlungsprozessen zu stärken.

Zu diesem Zweck wird in Kapitel 2 zunächst das zugrunde liegende soziotechnische Verständnis von KI erläutert und die Rolle von KI im Bildungssektor verortet. Es folgt eine Gegenüberstellung von Chancen und Risiken des KI-Einsatzes sowie eine kritische Diskussion gegenwärtiger Governance-Strategien. Darauf aufbauend wird das Argument für eine Refokussierung der öffentlichen Debatte und der Forschung auf eine kritischere Betrachtung der politischen und gesellschaftlichen Implikationen des KI-Einsatzes entwickelt und die Frage nach der Legitimation der KI-Einführung abgeleitet. Kapitel 3 entfaltet die Grundlagen der Legitimation politischer Entscheidungen und begründet ein pragmatisches Legitimitätsverständnis der Soziologie der Konventionen (Boltanski & Thévenot, 1991/2018; Diaz-Bone, 2018). Dieses Verständnis bildet den Ausgangspunkt für den konzeptionellen Rahmen, der in Kapitel 4 die Bedeutung der Politisierung und in Kapitel 5 die Rolle moralisch fundierter Rechtfertigung und

Kritik im Prozess der Herstellung legitimer politischer Entscheidungen bestimmt. Kapitel 6 erweitert diesen Rahmen um eine techniksoziologische Perspektive. Auf Basis der Rekonstruktion zentraler Deutungsmuster und historischer Technikerfahrungen wird gezeigt, wie diese Muster die Wahrnehmung und Bewertung moderner Technologien prägen. Diese theoretische und kontextuelle Rahmung bildet die Grundlage für die Ableitung der analyseleitenden Forschungsfragen in Kapitel 7. Sie richten sich auf die Muster der Rechtfertigung und Kritik, auf daraus entstehende Konfliktlinien sowie, vor dem Hintergrund eines zunehmend flächendeckenden KI-Einsatzes zur Lösung gesellschaftlicher Probleme, auf die Aufdeckung solutionistischer Argumentationsstränge.

Die empirische Untersuchung stützt sich auf eine inhaltlich strukturierende qualitative Inhaltsanalyse von KI-Policy-Dokumenten bildungspolitischer Akteur:innen. Berücksichtigt werden parlamentarische Dokumente auf Bundes- und Landesebene sowie Materialien weiterer Akteur:innen, die maßgeblichen Einfluss auf politische Entscheidungsprozesse zum Einsatz von KI im Bildungsbereich ausüben, darunter Beratungsgremien und Gewerkschaften. Kapitel 8 beschreibt das methodische Vorgehen, einschließlich der systematischen Auswahl der untersuchten Dokumente und der Durchführung der qualitativen Inhaltsanalyse. Die empirischen Ergebnisse werden in Kapitel 9 dargestellt und in Kapitel 10 im Lichte des konzeptuellen Rahmens kritisch diskutiert. Kapitel 11 reflektiert die zentralen Erkenntnisse der Arbeit aus der Perspektive der Wissenschaft, Politik und Zivilgesellschaft und gibt einen Ausblick auf weitere Forschungsbedarfe und Entwicklungen.

2. Künstliche Intelligenz als soziotechnisches System verstehen

„KI-Ära“ wurde kürzlich von der Gesellschaft für deutsche Sprache (2025) zum Wort des Jahres gewählt. So hat wohl keine andere technologische Entwicklung in den letzten Jahren eine vergleichbare Aufmerksamkeit auf sich gezogen. KI steht zugleich für das Versprechen von Effizienz, Innovation und Problemlösung sowie für Sorgen vor Überwachung, Kontrollverlust und gesellschaftlicher Disruption (Callaghan, 2018; Diaz-Bone & Schrör, 2022; Gallego & Kurer, 2022; Katzenbach & Ulbricht, 2019). Vor dem Hintergrund dieser polarisierten Wirkungszuschreibungen erfordert ein differenzierter Zugang zum Phänomen eine definitorische Bestimmung des KI-Begriffs, die sowohl technische Aspekte als auch die gesellschaftliche Einbettung berücksichtigt. Im Zentrum steht dabei das Verständnis von KI als soziotechnisches System, das technologische Verfahren in institutionelle, normative und kulturelle Kontexte eingebettet sieht (Kudina & van de Poel, 2024; Richter et al., 2023). Dieses Verständnis bildet die Grundlage für eine fundierte Auseinandersetzung mit Fragen der Legitimation, Regulierung und Konflikthaftigkeit, die im weiteren Verlauf der Arbeit vertiefend behandelt werden.

Historisch betrachtet ist KI dabei kein neues Phänomen. Bereits in den 1940er Jahren stellte Alan Turing die Frage, ob Maschinen denken können, und schlug mit dem Turing-Test ein erstes formales Kriterium maschineller Intelligenz vor (Haenlein & Kaplan, 2019). Der Begriff „Artificial Intelligence" wurde jedoch erstmalig 1956 im Rahmen der Dartmouth-Konferenz geprägt, die als Gründungsmoment des (heute) interdisziplinären Forschungsfelds der KI gilt (McCarthy et al., 1955). Die frühe Forschung orientierte sich stark an symbolischen Verfahren und Logik, mit dem Ziel, menschliches Denken zu formalisieren und in regelbasierte Systeme zu übersetzen (Haenlein & Kaplan, 2019). Während KI damit zunächst primär als wissenschaftlich-technisches Artefakt verhandelt wurde, fand sie 1984 erstmalig politische Berücksichtigung im Deutschen Bundestag (Willems, 2021). Erst seit den 2010er Jahren steigen jedoch die Verbreitung KI-basierter Systeme und die öffentliche Aufmerksamkeit zunehmend. Angetrieben wird dieser Aufschwung einerseits durch technischen Fortschritt, insbesondere durch den exponentiellen Zuwachs verfügbarer Daten (Big Data), Weiterentwicklungen im maschinellen Lernen (ML) sowie eine erhebliche Steigerung der Rechenkapazitäten (Duan et al., 2019; Hirsch-Kreinsen, 2024; O'Leary, 2013; Wischmann & Rohde, 2019), und andererseits durch neoliberale Gesellschaftsstrukturen (Gallego & Kurer, 2022; siehe Abschnitt 2.1).

Die begriffliche Verwendung von KI ist seit jeher mehrdeutig. In öffentlichen, wissenschaftlichen und politischen Debatten werden Begriffe wie ML, Deep Learning oder Algorithmus oft synonym gebraucht, obwohl sie unterschiedliche Konzepte bezeichnen. Während Algorithmen allgemein regelbasierte Verfahren darstellen, verweisen ML und Deep Learning auf spezifische Lernverfahren, welche aus Daten generalisieren und adaptiv reagieren und heute das Fundament vieler KI-Anwendungen bilden (Cormen et al., 2022; Janiesch et al., 2021; Kühl et al., 2022; Samoili et al., 2020). KI selbst wird häufig als Fähigkeit definiert, kognitive Leistungen zu erbringen, die typischerweise Menschen vorbehalten sind, wie logisches Denken, Lernen oder Problemlösen (Bartneck et al., 2021; Morandín-Ahuerma, 2022). Diese Orientierung verwischt jedoch die Grenze zwischen „schwacher" KI, die auf eng umrissene Aufgaben spezialisiert ist, und „starker" KI, die als hypothetische menschenähnliche Intelligenz mit einem hohen Maß an Autonomie und Flexibilität konzipiert wird, bisher aber nicht existiert (Fjelland, 2020; Jiang et al., 2022). Für die gegenwärtige Analyse ist daher der Fokus auf schwache KI sachgerecht, da diese die politische und gesellschaftliche Realität prägt.

Funktional definieren Haenlein und Kaplan (2019) KI als Systeme, die Daten interpretieren, aus ihnen lernen und auf sie reagieren, um vorgegebene Ziele zu erreichen. Die OECD (2019) ergänzt dies um eine prozesshafte Perspektive, die Wahrnehmung, Modellbildung und Inferenz betont und zugleich unterschiedliche Autonomiegrade berücksichtigt:

> An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. It does so by using machine and/or human-based inputs to: i) perceive real and/or virtual environments; ii) abstract such perceptions into models through analysis in an automated manner (e.g. with ML, or manually); and iii) use model inference to formulate options for information or action. AI systems are designed to operate with varying levels of autonomy. (S. 23 f.)

Anhand dieser Definition wird deutlich, dass KI kein einheitliches System bezeichnet, sondern ein Spektrum an Verfahren und Anwendungsgebieten mit unterschiedlichen Komplexitätsgraden umfasst. In der Praxis steht KI somit für eine Vielzahl von Anwendungen, die sich in ihrer Funktionsweise, Zielsetzung und gesellschaftlichen Relevanz teils erheblich unterscheiden (Krafft et al., 2022; Ulbricht, 2020), wodurch eine Pluralisierung des KI-Begriffsverständnisses zu beobachten ist. Diese definitorische Offenheit birgt zugleich praktische Konsequenzen: Insbesondere für Forschung und Technikregulierung stellt sich die Herausforderung, KI zu klassifizieren und zu gestalten, obwohl es an einer allgemein anerkannten Definition mangelt (Bhatnagar et al., 2018; Monett & Lewis, 2018; Wang, 2019). So führt die Mehrdeutigkeit der KI-Verständnisse zu unterschiedlichen Wahrnehmungen und Erwartungen ihrer möglichen Konsequenzen (Brynjolfsson et al., 2017; Kerr et al., 2020; Lünich et al., 2024a). Auf Grundlage dieser Wahrnehmungen entwickeln Akteur:innen divergente Interessen und (Policy-)Präferenzen (Bareis & Katzenbach, 2022; Cave & Dihal, 2019; Nokkala, 2016; Richter et al., 2023, 2025; Taeihagh, 2021), die unterschiedliche Rechtfertigungslogiken zur Legitimation ihres Handelns begründen können. Diese Differenzen bergen ein Konfliktpotenzial, welches insbesondere in der Debatte um Technikregulierung zum Ausdruck kommt und im Rahmen der *Science and Technology Studies* (STS) aufgegriffen wird (siehe Kapitel 6).

Angesichts dieser definitorischen Mehrdeutigkeit und der Herausforderung einer inkonsistenten Probleminterpretation wird daher deutlich, dass ein rein technisches Begriffsverständnis an dieser Stelle zu kurz greift. Im Einklang mit etablierten Positionen der STS (z. B. Jasanoff, 2016; Rammert, 1993) muss KI vielmehr als soziotechnisches System verstanden werden, welches erstens in gesellschaftliche Strukturen eingebettet und dabei zweitens nicht nur Mittel zur Zielerreichung, sondern Ausdruck kollektiver Interessen, Werte und Machtverhältnisse ist. Denn was als „intelligent“ gilt, ist nicht nur eine Frage algorithmischer Leistungsfähigkeit, sondern ebenso Ergebnis gesellschaftlicher Deutungsmuster. KI-Anwendungen sind in diesem Sinne keine neutralen Werkzeuge, sondern Ausdruck normativer Entscheidungen und gesellschaftlicher Rationalitäten, etwa bei der Auswahl und Gewichtung von Daten oder

Funktionszuschreibungen der KI (Bartl, 2023; Diaz-Bone & Schrör, 2022; Kudina & van de Poel, 2024; Mittelstadt et al., 2016; O'Neil, 2016).

Gerade deshalb erscheint es notwendig, KI im Kontext ihrer Einsatzgebiete sowie der gesellschaftlichen und politischen Strukturen zu betrachten, in die sie eingebettet ist. Denn die Folgen und gesellschaftlichen Implikationen von KI variieren erheblich je nach Funktionalität und Anwendungsfeld (Krafft et al., 2022; Ulbricht, 2020; Zweig, 2019): Eine algorithmische Empfehlung auf einer Unterhaltungsplattform ist weit weniger konfliktträchtig als eine automatisierte Entscheidung in migrations-, sicherheits- oder bildungspolitisch sensiblen Bereichen. Besonders im öffentlichen Sektor kommt KI-Anwendungen eine besondere Bedeutung zu, da Bürger:innen deren Auswirkungen nicht ohne Weiteres entgehen können und Fehlurteile mit einer hohen Tragweite verbunden sind.

2.1. Der Aufschwung Künstlicher Intelligenz im Bildungssektor als Ausdruck neoliberaler Restrukturierung

Aufgrund der definitorischen Mehrdeutigkeit des KI-Begriffs sowie der Bedeutung des soziotechnischen Kontexts erscheint es notwendig, die vorliegende Arbeit entlang eines konkreten Anwendungsfalls auszurichten. Sowohl die Wahrnehmungen von KI (Hampel et al., 2018; Kerr et al., 2020; Kropp & Hauser, 2024; Marcinkowski & Flaßhoff, 2024; Starke et al., 2022) als auch die Legitimation und Rechtfertigung politischer Entscheidungen können dabei je nach Anwendungskontext variieren (Forst, 2015; T. Johnson et al., 2024). Die Arbeit wählt daher das Bildungswesen als exemplarischen Anwendungsfall des KI-Einsatzes im öffentlichen Sektor. Diese Fokussierung dient der analytischen Eingrenzung, ohne die Übertragbarkeit der Erkenntnisse auf weitere Anwendungsfelder auszuschließen. Gerade weil sich Rechtfertigungsanforderungen aus der jeweiligen Einbettung von KI in gesellschaftliche Strukturen ergeben, ermöglicht ein konkreter Anwendungsfall jedoch, die Dynamiken von Rechtfertigung und Kritik präziser zu erfassen.

Wenngleich der Bildungssektor lange Zeit nur vereinzelt durch algorithmische Systeme unterstützt wurde (Attaran et al., 2018; Büching et al., 2019; Schiff, 2021), ist in den vergangenen Jahren ein deutlicher Anstieg des KI-Einsatzes zu verzeichnen (Holmes & Tuomi, 2022; Nemorin et al., 2023). Neben Anwendungen im Bereich des adaptiven und personalisierten Lernens sowie der Lehrunterstützung kommen KI-Systeme zunehmend auch in hochsensiblen Bereichen zum Einsatz, wie zum Beispiel bei der automatisierten Leistungsbewertung oder im Rahmen administrativer Prozesse wie der Studienzulassung (Chiu et al., 2023; Keller et al., 2019; Wang et al., 2024; Zawacki-Richter et al., 2019). Dem Bildungssektor kann damit ein besonderes Schadenspotenzial zugesprochen werden: Während auf KI-Entscheidungen basierende

Fehlurteile in privatwirtschaftlichen Kontexten bereits schwerwiegende Folgen haben können, sind sie im öffentlichen Sektor gerade deshalb besonders legitimationsbedürftig, da Betroffene sich den Entscheidungen in der Regel nicht entziehen können und das Ausmaß potenzieller Fehlentscheidungen besonders hoch bewertet werden muss (Krafft et al., 2022; Zweig, 2019). Der europäische *AI Act* trägt dieser Einschätzung Rechnung, indem er KI-Systeme im Bildungswesen als Hochrisikoanwendungen einstuft (Europäische Union, 2024; Lünich et al., 2024a). Aufgrund der Tragweite der Konsequenzen der KI-Anwendungen ist im Vergleich zu privatwirtschaftlichen Anwendungen daher in der öffentlichen Auseinandersetzung sowohl mit höherer Kritik in entsprechenden Anwendungsbereichen als auch mit höheren Anforderungen an den KI-Einsatz zu rechnen, was zugleich einen höheren Legitimationsbedarf und somit Rechtfertigungsdruck auslöst (Diaz-Bone & Schrör, 2022; Graß, 2015). Besonders deutlich wird dies am Beispiel des staatlichen französischen Universitätszulassungssystems *Parcoursup*, das 2018 breite Kritik auf sich zog, da die KI-basierten Entscheidungen soziale Ungleichheiten reproduzierten und intransparent blieben (Carpentier & Courtois, 2025; Frouillou et al., 2020). Der Fall zeigt, dass KI-basierte Entscheidungen im Bildungsbereich nicht nur technische, sondern zentrale politische und normative Fragen berühren, etwa nach individuellen Bildungschancen, demokratischer Kontrolle oder gesellschaftlicher Teilhabe. Damit bildet der Bildungssektor ein paradigmatisches Feld, in dem sich die Konfliktträchtigkeit algorithmischer Entscheidungsprozesse exemplarisch verdichtet.

Die Untersuchung des Bildungssektors ist zugleich vor dem Hintergrund übergeordneter gesellschaftlicher Transformationsprozesse zu verstehen. So lässt sich die Entwicklung des KI-Einsatzes nicht allein durch technische Fortschritte erklären, sondern sie wird durch neoliberale Restrukturierungen des öffentlichen Sektors verstärkt. Seit den 1980er Jahren haben sich mit dem neoliberalen Leitbild des *New Public Managements* Effizienz, Wettbewerb und Leistungssteigerung als zentrale Steuerungsmaximen im Bildungswesen etabliert (Graß, 2015; Hood, 1991; Mintz, 2021; Schimank, 2005; Turnbull et al., 2024). Bildungseinrichtungen reagieren dabei auf eine Reduktion staatlicher Finanzierung sowie auf die Forderung eines nachweislichen gesellschaftlichen Nutzens von Forschung und Lehre mit einer stärkeren Ausrichtung auf Wettbewerb und Messbarkeit (Lange & Schimank, 2007). Der künstlich erzeugte Konkurrenzdruck soll die Leistungsfähigkeit und Effizienz der Institutionen erhöhen, führt jedoch zugleich zu neuen Abhängigkeiten von externen Evaluations- und Steuerungsmechanismen (Bleiklie, 2018; Lane, 2000).

In diesem Kontext gewinnen zum einen Formen der Metrisierung und Automatisierung an Bedeutung. Metrisierung bezeichnet dabei den Versuch, Lernverhalten und -erfolg oder auch

wissenschaftliche Outputs zu standardisieren und in messbare Kennziffern und Rankings zu überführen (Krücken, 2017; Santori, 2023). Diese Entwicklung fügt sich in eine breitere gesellschaftliche Tendenz ein, die Fourcade und Healy (2024) in einer Gegenwartsdiagnose als zentralen Steuerungsmechanismus gegenwärtiger Gesellschaften identifizieren: „We live in an ordinal society, a society oriented toward, justified by, and governed through measurement" (S. 1). Vor diesem Hintergrund erscheinen KI-Systeme als logische Fortsetzung dieser neoliberalen Entwicklungen. Sie bauen auf der zunehmenden Verfügbarkeit quantifizierter Daten auf, automatisieren deren Auswertung und versprechen, Entscheidungen effizienter, standardisierter und objektivierter zu gestalten. Damit knüpfen sie unmittelbar an bestehende Steuerungsrationalitäten an und verstärken die bereits etablierte Orientierung an Messbarkeit und Vergleichbarkeit im Bildungswesen (Ball & Grimaldi, 2022; Saltman, 2020).

Zum anderen eröffnet diese Entwicklung Raum für den Einfluss privater Unternehmen, die diesen leistungsorientierten Reformprozess nutzen, um Bildungstechnologien zu vermarkten, die etwa höhere Studienabschlussquoten versprechen (Castañeda & Selwyn, 2018; Komljenovic et al., 2025; Smithers, 2020). Kritiker:innen sehen darin eine Vertiefung neoliberaler Steuerungsrationalitäten, die Bildung zunehmend unter ökonomische Zielvorgaben stellt und soziale Gerechtigkeitsfragen ebenso in den Hintergrund drängt wie demokratische Kontrolle (Ball & Grimaldi, 2022; Mintz, 2021; Sahlberg, 2023; Saltman, 2020; Selwyn, 2024; Slade & Prinsloo, 2013). Damit birgt die neoliberale Restrukturierung des Bildungssektors nicht nur erhebliche Konfliktpotenziale, sondern fördert zugleich eine Politisierung handlungsleitender Werte der Bildungspolitik, die durch den KI-Einsatz herausgefordert werden (Turnbull et al., 2024).

Das soziotechnische Verständnis von KI im Bildungssektor macht unmittelbar auf die Ambivalenz technischer Innovationen aufmerksam. Einerseits stehen Versprechen von Effizienzsteigerung und Wettbewerbsfähigkeit im Raum, andererseits treten Risiken von Diskriminierung, Fehlurteilen und einer Erosion demokratischer Teilhabe hervor, die auf ein Konfliktpotenzial verweisen. Um die Legitimität der Einführung von KI-Systemen im Bildungsbereich und darüber hinaus im öffentlichen Sektor insgesamt angemessen zu erfassen, ist es daher notwendig, sowohl die Potenziale als auch die Risiken dieser Technologie zu beleuchten. Erst durch die Gegenüberstellung dieser Dimensionen wird der besondere Legitimationsbedarf sichtbar: Risiken markieren zentrale Bezugspunkte der Kritik, die einen Rechtfertigungsdruck aufrechterhält, während Potenziale den positiven Bezugspunkt liefern, an dem sich Rechtfertigungen orientieren können.

2.2. Potenziale Künstlicher Intelligenz für Gesellschaft und Bildung

Die Einführung moderner Technologien ist historisch betrachtet stets mit einer gewissen Ambivalenz zwischen dem Versprechen gesellschaftlicher Verbesserungen und der Sorge vor unerwünschten Nebenfolgen verbunden. Diese Ambivalenz wird im Fall der KI besonders deutlich, da KI nicht nur als technische Innovation, sondern als potenziell tiefgreifende Transformationskraft nahezu aller gesellschaftlichen Lebensbereiche wahrgenommen wird (Moravec & Martínez-Bravo, 2023; Păvăloaia & Necula, 2023). In zahlreichen Anwendungsfeldern, von Medizin über Industrie bis hin zur Krisenbewältigung, werden mit KI Hoffnungen auf Effizienzsteigerung, Präzision und Individualisierung verknüpft (Cools et al., 2024; Kim et al., 2022; Malik et al., 2021). Ihr zentraler Effekt liegt dabei in der datenbasierten Optimierung komplexer Entscheidungs- und Handlungsprozesse. Auch im Bildungssektor wird KI zunehmend als Schlüsseltechnologie verstanden, der ein breites Spektrum an Potenzialen zur Weiterentwicklung und Modernisierung von Lehr-, Lern- und Steuerungsprozessen zugeschrieben wird (Attaran et al., 2018; Chiu et al., 2023; Zawacki-Richter et al., 2019). Während die KI-Anwendungsgebiete variieren, lassen sich die ihnen zugeschriebenen Potenziale entlang fünf Dimensionen kategorisieren: (1) effizienzorientierte Optimierung und (2) Objektivierung bestehender Prozesse, (3) Adaptivität und Individualisierung sowie (4) demokratiefördernde und schließlich (5) kreative Potenziale.

Erstens wird ein hohes Potenzial in der Prozessoptimierung durch KI-Anwendungen gesehen. KI soll helfen, komplexe Aufgaben mit geringeren Mitteleinsätzen zu bewältigen, indem sie große Datenmengen analysiert, Muster erkennt, Prognosen erstellt und Entscheidungen vorbereitet oder trifft (Baabdullah, 2024; M. Boyd & Wilson, 2017; Elhaddad & Hamam, 2024; Engin & Treleaven, 2019; Mohanasundari et al., 2023; O'Reilly, 2013; Umoga et al., 2024; Wirtz & Müller, 2019). Die Automatisierung repetitiver Aufgaben gilt dabei sowohl in Wirtschaft und Verwaltung als auch im Bildungswesen als Schlüssel zur Effizienzsteigerung und zu Rationalisierungsgewinnen. Dabei sollen KI-Systeme etwa Prüfungsprozesse, Studienzulassungen oder die Stundenplanung unterstützen. Durch die Analyse großer, heterogener Datenbestände verspricht KI darüber hinaus eine Reduktion von Fehlern und eine optimierte Ressourcennutzung (Eden et al., 2024; Labadze et al., 2023; Lampou, 2023; Tsai et al., 2020).

Mit diesen funktionalen Effekten verknüpft ist zweitens das Potenzial der Objektivierung von Entscheidungen. Im Unterschied zu menschlichen Entscheidungen soll KI auf Grundlage datenbasierter Muster urteilen, ohne von emotionalen, kognitiven oder sozialen Verzerrungen beeinflusst zu sein (Lee, 2020; Shrestha et al., 2019; Simmons, 2018; Steyvers & Kumar, 2024). Dieses Objektivitätsversprechen knüpft insbesondere im Bildungsbereich an die Hoffnung,

Bewertungsprozesse, Zulassungsentscheidungen und Leistungsdiagnosen fairer gestalten zu können (Apetorgbor et al., 2024; Chai et al., 2024; Sato et al., 2024).

Drittens kommen der Adaptivität und Individualisierung zentrale Rollen zu. Basierend auf der Fähigkeit zur Mustererkennung wird KI in zahlreichen Anwendungsbereichen dafür eingesetzt, Prozesse dynamisch an individuelle Bedürfnisse anzupassen. So werden Vorhersagen über zukünftige Entwicklungen getroffen und Empfehlungen oder Entscheidungen situativ entlang der analysierten Muster angepasst. Solche individualisierten Lösungen gelten als Beitrag zur Anpassung von Nutzendenbedürfnissen und Angebotsstruktur (Dwivedi et al., 2021). Im Bildungsbereich wird dieses Potenzial besonders betont: KI-basierte Systeme wie Learning Analytics oder intelligente Tutorensysteme analysieren Lernverläufe und generieren darauf abgestimmt individualisiertes Feedback, Lernempfehlungen und Unterstützungsmaßnahmen. Diese Anwendungen sollen Motivation und Lernerfolg steigern und langfristig zu höheren Abschlussquoten beitragen (Alturki et al., 2022; Attaran et al., 2018; Chiu et al., 2023; Lünich et al., 2024b; Pataranutaporn et al., 2021; Zhang & Aslan, 2021).

Viertens werden KI demokratiefördernde Potenziale zugeschrieben. So können KI-Systeme etwa zur Unterstützung öffentlicher politischer Diskurse beitragen, indem sie große Informationsmengen analysieren und Stimmungen und Konfliktlinien frühzeitig erkennen. Auf dieser Basis lassen sich Entscheidungsprozesse dateninformiert unterstützen (Boella et al., 2016; König & Wenzelburger, 2020; Marmolejo-Ramos et al., 2022; Savaget et al., 2018). Darüber hinaus ermöglichen Chatbots oder partizipative Plattformen, die auf KI basieren, eine niedrigschwellige politische Teilhabe und können neue Formen des gesellschaftlichen Dialogs eröffnen. Diesem Inklusionspotenzial wird insbesondere im Bildungsbereich hohe Bedeutung zugemessen. Hier sollen beispielsweise durch KI-basierte Übersetzungen Lernbarrieren reduziert und Chancengerechtigkeit gefördert werden (Labadze et al., 2023; Melo-López et al., 2025; Reyes & Meneses, 2024; Song et al., 2024; Viberg et al., 2024).

Über diese normativen Potenziale hinaus wird KI schließlich fünftens ein kreatives und exploratives Potenzial zugeschrieben, das über klassische Automatisierung hinausgeht. Sie kann nicht nur Antworten auf bekannte Fragen liefern, sondern auch neue Fragestellungen generieren. Überraschungen gelten dabei nicht als Fehlfunktion, sondern als konstitutives Merkmal der Technologie, so Anicker (2023): „Anders als bei Rasenmähern und Waschmaschinen sind Überraschungen der Nutzer in diesem Fall nicht *‚bug‘* sondern *‚feature‘*. Von KI erwarten wir Antworten auf Fragen, die wir nicht gestellt haben“ (S. 141; Herv. i. O.). Diese epistemische Offenheit macht KI zu einer Quelle der Wissensproduktion, der insbesondere im Kontext

generativer KI ein hohes Kreativitätspotenzial zugesprochen wird (Chen et al., 2023; Creely & Blannin, 2025). Somit soll KI Prozesse beschleunigen, Entscheidungen verbessern, Bedürfnisse individueller adressieren, demokratische Teilhabe fördern und neue Erkenntnisse ermöglichen. Doch gerade die Ausmaße der Potenziale erklären, warum KI zugleich als Projektionsfläche gesellschaftlicher Ängste fungiert. So steigen mit wachsender Reichweite und Wirksamkeit auch die Anforderungen an die Regulierung und politische Steuerung von KI. Dies wird vor allem in der Diskussion um die mit der Technologie verbundenen Risiken sichtbar.

2.3. Risiken Künstlicher Intelligenz für Gesellschaft und Bildung

Die Risiken, die mit dem KI-Einsatz verbunden sind, erweisen sich als ebenso vielfältig und weitreichend wie die ihr zugeschriebenen Potenziale. Wirtz et al. (2022) systematisieren basierend auf einer umfassenden Literaturanalyse sechs zentrale Risikokategorien. Diese umfassen (1) technologische, (2) informationelle, (3) ökonomische, (4) soziale und gesellschaftliche, (5) ethische sowie (6) rechtlich-regulatorische Risiken. Diese Einteilung dient im Folgenden einer differenzierten Betrachtung möglicher negativer Effekte KI-basierter Systeme.

Technologische Risiken von KI-Systemen umfassen neben sicherheitsbezogenen Gefahren, zum Beispiel der Manipulation durch Hackerangriffe, auch die Gefahr fehlerhafter Entscheidungen infolge von Programmierfehlern oder unzureichender Datenqualität (M. Boyd & Wilson, 2017; Danaher et al., 2017; Dwivedi et al., 2021; Wirtz et al., 2019, 2022). Darüber hinaus liegt ein zentrales technisches Risiko in der Komplexität der KI-Systeme begründet. Dabei erschweren selbstlernende Verfahren die Nachvollziehbarkeit des Systemverhaltens, was insbesondere bei autonomen Entscheidungen ohne menschliches Eingreifen zu Kontrollverlusten führen kann (M. Boyd & Wilson, 2017; Burrell & Fourcade, 2021; Scherer, 2016). Komplexe ML-Verfahren, die beispielsweise die Grundlage für automatisierte Leistungsbewertungen bilden können, operieren häufig als sogenannte „Black Boxes“, deren Entscheidungslogik weder für Entwickler:innen noch für Anwender:innen vollständig nachvollziehbar ist (Keller et al., 2022; Prinsloo, 2020). Mangelnde Transparenz beispielsweise darüber, auf welchen Daten, Gewichtungen oder Entscheidungsregeln die Bewertung beruht, untergräbt die Möglichkeit, algorithmische Urteile kritisch zu prüfen oder anzufechten (Burrell, 2016; Lepri et al., 2018; Shin & Park, 2019; Warner & Sloan, 2021). Damit beeinflusst sie nicht nur die Wahrnehmung der Fairness und Angemessenheit solcher Systeme, sondern auch deren Legitimation in sensiblen Bildungsprozessen (de Bruijn et al., 2022; K. de Fine Licht & de Fine Licht, 2020; Langer & König, 2023; Lünich & Keller, 2024; Lünich et al., 2024a; Pasquale, 2015; Williams et al., 2022).

Informationelle Risiken des KI-Einsatzes betreffen weniger spezifische Anwendungen im Bildungssystem selbst, sondern wirken vor allem über die Medien- und Informationslandschaft auf Bildungs- und Meinungsbildungsprozesse ein. Generative KI-Systeme können täuschend echte Inhalte erzeugen und damit zur Verbreitung von Desinformation beitragen. Deepfakes oder automatisierte Fake-News-Kampagnen untergraben die Verlässlichkeit öffentlicher Wissensräume und erschweren die Unterscheidung zwischen glaubwürdigen und manipulierten Informationen (Chesney & Citron, 2019; König & Wenzelburger, 2020; Pawelec, 2022; van der Sloot & Wagensveld, 2022; Westerlund, 2019). Zugleich können algorithmische Filter- und Moderationssysteme dazu führen, dass Inhalte automatisiert gelöscht oder in ihrer Reichweite beschränkt werden, was mit dem Risiko (unbeabsichtigter) Zensur und eingeschränkter Pluralität einhergeht (Just & Latzer, 2017; König & Wenzelburger, 2020; Rahwan, 2018). Damit betreffen informationelle Risiken zentrale Voraussetzungen demokratischer Bildung: den freien Zugang zu verlässlichen Informationen, die Fähigkeit zur kritischen Urteilsbildung und die Offenheit öffentlicher Diskurse (Jungherr, 2023; Thierer et al., 2017; Wirtz & Müller, 2019; Wirtz et al., 2022).

Ökonomische Risiken des KI-Einsatzes entstehen vor allem durch Automatisierungsprozesse, die Arbeitsmärkte und Qualifikationsanforderungen tiefgreifend verändern. Der zunehmende Einsatz KI-gestützter Systeme führt insbesondere in standardisierten Tätigkeitsfeldern zu Beschäftigungsverlusten, während neue, hochspezialisierte Qualifikationen entstehen. Diese Verschiebungen bergen Risiken sozialer Ungleichheit und struktureller Erwerbslosigkeit, die sich auch im Bildungsbereich etwa durch veränderte Anforderungen an Lehrende oder eine stärkere Ökonomisierung von Bildung niederschlagen können (M. Boyd & Wilson, 2017; Du, 2024; Korinek & Stiglitz, 2019; Scherer, 2016; Selwyn, 2019). Damit stellt sich nicht nur die Frage nach dem Umgang mit disruptiven Beschäftigungseffekten, sondern auch nach einer gerechten Verteilung der Produktivitätsgewinne, die KI-Systeme erzeugen (Korinek & Stiglitz, 2019).

Mit den ökonomischen Folgen der Automatisierung gehen darüber hinaus soziale Risiken einher. So können Arbeitsplatzverluste und veränderte Qualifikationsanforderungen gesellschaftliche Spannungen und Ungleichheiten verstärken (Thierer et al., 2017; Winfield & Jirotka, 2018; Wirtz et al., 2022). Gleichzeitig wirft der Einsatz von KI-Systemen grundlegende Fragen des Datenschutzes und der Privatsphäre auf. Viele Anwendungen verarbeiten personenbezogene Daten, um Lernverhalten, Leistungen oder administrative Entscheidungen zu analysieren, häufig ohne ausreichende Transparenz über Zweck und Reichweite der Datennutzung (boyd & Crawford, 2012; Rubel & Jones, 2016). Intransparente Datensammlungen oder die Kombination verschiedener Datenquellen können dabei zu schwer kontrollierbaren Profilen führen und

die Autonomie des Einzelnen einschränken (Dogruel et al., 2022; Laitinen & Sahlgren, 2021; Manheim & Kaplan, 2019; Taylor, 2017; Yeung, 2018). Vor dem Hintergrund wachsender Überwachungs- und Kontrollmöglichkeiten steht damit nicht zuletzt das Recht auf informationelle Selbstbestimmung als zentrales Schutzgut demokratischer Gesellschaften zur Disposition (Williamson & Prybutok, 2024). Lernende äußern entsprechend Vorbehalte gegenüber der Freigabe persönlicher Daten zur KI-basierten Verarbeitung (Ifenthaler & Schumacher, 2016b; Lünich et al., 2024a).

Ethische Risiken des KI-Einsatzes werden Wirtz et al. (2022) zufolge insbesondere durch zwei Elemente bestimmt: Zum einen mangelt es vielen KI-Systemen an einer legitimen Grundlage, etwa in Hinblick auf ihre Entwicklung oder ihren weitreichenden Einsatz (Beckman et al., 2024; Wirtz & Müller, 2019). Zum anderen besteht trotz des Objektivitätsversprechens ein erhebliches Risiko der (Re-)Produktion von Diskriminierung (Barocas & Selbst, 2016; Kordzadeh & Ghasemaghaei, 2021; Žliobaitė, 2017). Diskriminierung kann sowohl aus der technischen Architektur als auch aus den verwendeten Trainingsdaten oder der Anwendungspraxis resultieren (Baker & Hawn, 2022; Fazelpour & Danks, 2021; Kizilcec & Lee, 2022; Mitchell et al., 2021). Statistische Verzerrungen entstehen etwa, wenn sensible Merkmale wie ethnische Herkunft oder Geschlecht direkt oder indirekt in algorithmische Entscheidungsprozesse einfließen. Inadäquate oder fehlende Fairnessmetriken in der technischen Entwicklung können so soziale Ungleichheiten produzieren (Kizilcec & Lee, 2022; Kleinberg et al., 2018; Makhlouf et al., 2021; Veale & Binns, 2017; Weyerer & Langer, 2019). Werden gesellschaftlich verankerte Vorurteile oder strukturelle Ungleichheiten hingegen in Trainingsdaten selbst eingeschrieben, reproduzieren Algorithmen diese in ihrem Output, auch ohne dass eine diskriminierende Absicht vorliegt (Broussard, 2018; Burrell & Fourcade, 2021; Eubanks, 2018; Fazelpour & Danks, 2021; Lepri et al., 2018; O'Neil, 2016). Doch Diskriminierung kann auch jenseits der Technik entstehen, etwa wenn KI-basierte Vorhersagen, Entscheidungen oder Empfehlungen von menschlichen Akteur:innen fehlinterpretiert oder unreflektiert übernommen werden. So kann ein vermeintlich objektives KI-System auch durch seine soziale Anwendungspraxis diskriminierende Effekte entfalten (Kizilcec & Lee, 2022; Lünich et al., 2024b). Besonders problematisch sind unbeabsichtigte Folgen wie selbsterfüllende Prophezeiungen: Wird etwa von einem KI-System prognostiziert, dass ein:e Schüler:in den Abschluss nicht erreichen wird, kann dies zu Demotivation und tatsächlichem Leistungsabfall sowie in der Konsequenz zum tatsächlichen Nichtbestehen führen (Keller et al., 2025; Madaio et al., 2022; Zeide, 2017). Dieses dreifache Diskriminierungspotenzial (im Modell, in den Daten und in der Nutzung) unterstreicht die Notwendigkeit eines ganzheitlichen, interdisziplinären Blicks auf die gesellschaftlichen Folgen und die

Legitimation des KI-Einsatzes. Zahlreiche Beispiele aus Bildung (Edwards, 2021; Frouillou et al., 2020; Smith, 2020; Wenzelburger & Hartmann, 2022), Justiz (Angwin et al., 2016; Simmons, 2018), Personalwesen (Köchling & Wehner, 2020) und Sozialpolitik (Allhutter et al., 2020; Bekker, 2021; Kaun, 2021) belegen, dass algorithmische Diskriminierung erhebliche Proteste hervorrufen kann, die die Legitimität von KI-Anwendungen nachhaltig infrage stellen.

Schließlich besteht ein zentrales rechtliches und regulatorisches Risiko von KI-Anwendungen in der mangelnden Verantwortlichkeit (*Accountability*) KI-basierter Entscheidungen (de Laat, 2018; Scherer, 2016; Shah, 2018). Je stärker KI-Systeme autonom agieren und je weniger ihre Entscheidungsprozesse nachvollziehbar sind, desto schwieriger wird es, im Falle von Fehlern oder Schäden klare Verantwortlichkeiten zu benennen (Busuioc, 2021; Diakopoulos, 2016; Williams et al., 2022). Die Vielzahl beteiligter Akteursgruppen, von Entwickler:innen über Betreiber:innen bis hin zu Anwender:innen, führt häufig zu einer Verantwortungsdiffusion, die Rechenschaftspflichten erschwert (Slota et al., 2023). In der Folge können Fehlentscheidungen ohne institutionelle Konsequenzen bleiben und so nicht nur individuelle Rechte, sondern auch das gesellschaftliche Vertrauen in politische Entscheider:innen oder KI-einführende Institutionen untergraben werden. Gleichzeitig liegt ein weiteres Risiko in der bislang unzureichenden rechtlichen Regulierung von KI-Systemen, die angesichts ihrer potenziellen Reichweite und Wirkung zunehmend erforderlich wird (Gasser & Almeida, 2017; Thierer et al., 2017; Wirtz et al., 2022).

Vor diesem Hintergrund stellt sich die Frage, wie ein sozialverantwortlicher, legitimer KI-Einsatz gesichert werden kann. Aktuelle Ansätze der Technikregulierung setzen dabei auf eine menschenzentrierte KI-Entwicklung, die auf eine Nivellierung der Potenziale und Risiken abzielt. Im Folgenden werden diese Bestrebungen skizziert und es wird argumentiert, dass sie zur Sicherung der Legitimität von KI bislang nur begrenzt beitragen, weshalb eine Refokussierung von Forschung und Regulierung erforderlich erscheint.

2.4. Das Versprechen menschenzentrierter KI-Entwicklung zur Sicherung legitimer Künstlicher Intelligenz

Mit der zunehmenden Verbreitung und Relevanz von KI-Systemen wächst nicht nur das Bewusstsein für deren transformatives und innovationsförderndes Potenzial, sondern auch für die damit verbundenen Risiken. Entsprechend werden Forderungen nach einer umfassenden Technikregulierung laut (Guihot et al., 2017; Haenlein & Kaplan, 2019; Krafft et al., 2022; Perry & Uuk, 2019; Taeihagh, 2021; Waldman & Martin, 2022; Züger & Asghari, 2023). Unter dem

Konzept der *Algorithmic Governance*[1] werden dabei Ansätze gefasst, die darauf abzielen, das Potenzial von KI-Systemen für Innovation und gesellschaftlichen Fortschritt nutzbar zu machen, ohne demokratische, rechtsstaatliche und gemeinwohlorientierte Prinzipien zu gefährden und die mit dem Einsatz verbundenen Risiken aus dem Blick zu verlieren (Hirsch-Kreinsen & Krokowski, 2024b; Justo-Hanani, 2022; Mäntymäki et al., 2022; Taeihagh, 2021).

Zentrale rechtliche Säulen dieser Regulierung bilden in Deutschland und auf europäischer Ebene vor allem die Datenschutz-Grundverordnung (DSGVO, 2016) sowie der EU *AI Act* (Europäische Union, 2024). Die DSGVO, seit 2018 in Kraft, setzt international anerkannte Datenschutzstandards und schreibt unter anderem Transparenzpflichten, Zweckbindung, Datenminimierung sowie Rechte auf Auskunft und Löschung fest (Albrecht, 2016; Williamson & Prybutok, 2024). Der *AI Act* verfolgt einen risikobasierten Regulierungsansatz, der KI-Systeme je nach Gefährdungspotenzial in Risikoklassen einteilt und für Hochrisikoanwendungen (z. B. in Strafverfolgung, Gesundheitswesen oder Bildung) strengere Anforderungen an Transparenz, Sicherheit, Dokumentation und menschliche Aufsicht vorsieht. Ziel ist es, ein einheitliches Regelwerk im europäischen Binnenmarkt zu etablieren, das Innovation fördert und zugleich grundlegende Rechte schützt (Novelli et al., 2025). Kritisch angemerkt wird jedoch, dass etwa unklare Begriffsbestimmungen (etwa der Risikokategorien) sowie Durchsetzungs- und Kontrollprobleme die Wirksamkeit des EU *AI Acts* einschränken (Hacker, 2024; Novelli et al., 2024; Veale & Zuiderveen Borgesius, 2021). Ebenso steht die DSGVO in der Kritik, dabei jedoch eher aus einer innovationshemmenden Perspektive: So sind ihre komplexen Anforderungen teils nur schwer mit den datenintensiven, adaptiven Prozessen vieler KI-Systeme vereinbar und drohen daher, insbesondere für kleinere Unternehmen, Innovationsprozesse zu hemmen (Geradin et al., 2021; Hacker et al., 2023; Hirsch-Kreinsen, 2023). Dabei besteht für die rechtlich bindenden Regularien die Herausforderung, konkrete Rahmenbedingungen für eine nahezu unbegrenzte Vielfalt möglicher Einsatzszenarien zu formulieren, einschließlich solcher, die zum Zeitpunkt ihres Inkrafttretens bisher nicht absehbar sind (Novelli et al., 2025).

Neben verbindlichen Rechtsnormen gewinnen Ethikleitlinien zunehmend an Bedeutung. Diese wurden in den vergangenen Jahren von internationalen Organisationen, Regierungen, Unternehmen sowie zivilgesellschaftlichen und akademischen Akteur:innen entwickelt und fließen

[1] Der Begriff der *Algorithmic Governance* wird in der wissenschaftlichen Debatte in zwei Bedeutungen verwendet (Cantero Gamito & Ebers, 2021): Erstens bezeichnet er die Regulierung algorithmischer Systeme im Sinne einer *Governance von Algorithmen*, die rechtliche und ethische Rahmenbedingungen für einen verantwortungsvollen KI-Einsatz entwickelt (Gasser & Almeida, 2017; Larsson, 2020; Mäntymäki et al., 2022; Taeihagh, 2021; Ulnicane et al., 2021). Zweitens kann er die Nutzung von KI-Anwendungen zur Steuerung politischer oder administrativer Abläufe im Sinne einer *Governance durch Algorithmen* meinen (Just & Latzer, 2017; Katzenbach & Ulbricht, 2019; Musiani, 2013; Sætra, 2020). In dieser Arbeit wird erstere Bedeutung zugrunde gelegt.

teilweise in rechtliche Regelwerke ein (Larsson, 2020). Ihr Ziel ist die Förderung einer menschenzentrierten und vertrauenswürdigen KI-Entwicklung, die individuelle Rechte achtet, gesellschaftlichen Nutzen stiftet und demokratische Prinzipien wahrt (Cantero Gamito & Ebers, 2021; Europäische Kommission, 2019; Lee et al., 2017; Shneiderman, 2020). Diese Leitlinien sollen normative Orientierung für die technische Entwicklung, Anwendung und politische Regulierung von KI bieten. Sie reagieren damit auf die in Forschung und Praxis identifizierten Risiken KI-basierter Systeme (siehe Abschnitt 2.3) und versuchen, diesen durch handlungsleitende ethische Prinzipien entgegenzuwirken.

Vergleichende Studien belegen einen weitgehenden internationalen Konsens bezüglich der Definition der ethischen Prinzipien. Diese umfassen Forderungen nach Transparenz und Erklärbarkeit, Fairness, Verantwortlichkeit, Datenschutz, technischer Sicherheit sowie der Wahrung von Autonomie und dem Schutz vor Schaden (Corrêa et al., 2023; Fjeld et al., 2020; Hagendorff, 2020; Jobin et al., 2019; Piano, 2020). Diese Prinzipien sollen sicherstellen, dass algorithmische Entscheidungen dem Gemeinwohl dienlich sind. So wird etwa durch Transparenz und Erklärbarkeit angestrebt, die Nachvollziehbarkeit algorithmischer Entscheidungen zu ermöglichen. Datenschutz soll zugleich individuelle Freiheitsrechte sichern, während Prinzipien wie Schadensverhütung (*Non-Maleficence*) oder Gemeinnützigkeit (*Beneficence*) darauf abzielen, technische Risiken zu vermeiden und gesellschaftlichen Nutzen zu stiften (Europäische Kommission, 2019; OECD, 2024).

Diese Governance-Bestrebungen stehen jedoch vor der Herausforderung, Dynamik und Komplexität der Technologie mit normativen Anforderungen in Einklang zu bringen. So stößt die konkrete Ausgestaltung der ethischen Prinzipien in der Praxis auf erhebliche Umsetzungsprobleme, da sie häufig in Zielkonflikten, sogenannten *Trade-offs*, zueinanderstehen (Bleher & Braun, 2023; Hagendorff, 2020; Piano, 2020; Whittlestone et al., 2019). Dabei kann etwa das Streben nach maximaler Vorhersagegenauigkeit (*Accuracy*) mit Datenschutzbestimmungen oder Transparenzansprüchen kollidieren (Barredo Arrieta et al., 2020; Felzmann et al., 2019; Holm, 2019; Petkovic, 2023; Sanderson et al., 2024). Die konkrete Gestaltung einer ethischen KI erfordert daher eine sorgfältige Aushandlung zwischen konkurrierenden Prinzipien und Interessen.

Zudem bleiben viele der geforderten ethischen Prinzipien in ihrer Beschreibung zu vage, um spezifische Risiken gezielt zu adressieren. So entsteht die Gefahr einer Ethikalisierung ohne bindende moralische Fundierung, wenn normative Grundsatzentscheidungen der KI-Entwicklung weitgehend Entwickler:innen überlassen werden, obwohl diese wertgebunden sind und im

Sinne einer demokratischen Legitimation einer öffentlichen Aushandlung bedürfen (Beckman et al., 2024; Bleher & Braun, 2023; Guihot et al., 2017; Larsson, 2020; Mittelstadt, 2019; Munn, 2023; Williams et al., 2022). Vor diesem Hintergrund weist Mittelstadt (2019) darauf hin, dass der vermeintliche Konsens über die Prinzipien über tiefgreifendere politische und normative Unstimmigkeiten hinwegtäusche. Weitere Kritiker:innen sprechen gar von „toothless principles", denen es an Spezifität, Verbindlichkeit und Konsequenzen mangele (Munn, 2023; Rességuier & Rodrigues, 2020). So hat sich bislang keine einheitliche Implementierungs- oder Governance-Strategie herauskristallisiert, die die Umsetzung der ethischen Prinzipien tatsächlich sichern könnte (Züger & Asghari, 2023). Der überwiegend nicht bindende Charakter vieler Leitlinien verstärkt dieses Problem, ebenso wie die Schwierigkeit, einheitliche Standards für unterschiedliche Anwendungsfelder zu formulieren (Butcher & Beridze, 2019; Guihot et al., 2017; Hagendorff, 2020; Krafft et al., 2022; Taeihagh, 2021). Ihre Eignung als effektives Instrument der KI-Regulierung wird daher zunehmend angezweifelt:

> At best, ethics guidelines are generically performative, operating at a linguistic level to assuage and deflect critique and regulation. Indeed, we argue that ethics discourses and solutions are currently operating as *assurance* for investors and the general public, rather than as an effective tool for governance (…). (Kerr et al., 2020, 10; Herv. i. O.)

Vor diesem Hintergrund plädieren einige Autor:innen für einen Ansatz der *Algorithmic Governance*, in dem die Sicherung der Legitimität der KI-Entwicklung nicht primär als Verantwortung der Entwickler:innen betrachtet wird oder durch technische Standards abzusichern ist, sondern durch einen demokratischen Aushandlungsprozess, in dem gesellschaftlich entschieden wird, welche KI-Anwendungen mit welchen Zielen und unter welchen Bedingungen als gerechtfertigt gelten (Beckman et al., 2024; Danaher et al., 2017; Katzenbach & Ulbricht, 2019; Züger & Asghari, 2023). So fordern Beckman et al. (2024) ebenso wie Züger und Asghari (2023), nicht allein die Rechtmäßigkeit einzelner von KI-Systemen getroffener Entscheidungen, sondern die Frage der Legitimität des gesamten KI-Einsatzes in den Vordergrund zu stellen. Im Zentrum stehe die Frage der Rechtfertigung des Systems, welches in erster Linie dem öffentlichen Interesse im Sinne des Allgemeinwohls dienen müsse:

> For AI to actually serve the public interest, we believe that first of all a justification to the public is necessary, to argue why the technology is not developed for the mere sake of innovation or commercial benefits but to serve a common public interest. The entity considering an AI system to be part of a solution in a social, policy or other public context, needs to present an argument to fellow citizens, giving the reasons how the system will

> tackle and improve the given issue and why it's the best solution in consideration of the alternatives. The reasons given need to be based on the democratic argument of the public concerned (which might be rights formulated in a constitution or other laws or other socially agreed on common goals). (Züger & Asghari, 2023, S. 819)

Dieser Perspektivwechsel verweist auf zwei zentrale Voraussetzungen für die Legitimation der KI-Einführung. Erstens bedarf es einer Politisierung von KI, die den KI-Einsatz nicht als rein technisch-administrative Frage behandelt, sondern als Gegenstand öffentlicher Aushandlung über gesellschaftliche Ziele. Nur so kann die normative Fundierung des soziotechnischen Systems demokratisch bestimmt werden. Zweitens ist eine öffentliche Rechtfertigung des KI-Einsatzes erforderlich, die plausibel macht, inwiefern dieser dem Gemeinwohl dient und alternative Lösungen berücksichtigt. *Algorithmic Governance* muss daher nicht nur normativ sensibel, sondern auch demokratisch verfasst sein, wenn sie der gesellschaftlichen Tragweite algorithmischer Systeme gerecht werden und ihre Einführung öffentlich legitimieren will (Beckman et al., 2024). Im nachfolgenden Abschnitt wird dieses Verständnis weiter ausgeführt und darauf aufbauend die forschungsleitende Fragestellung der Arbeit abgeleitet.

2.5. Forderung einer Refokussierung der Forschung und politischen Steuerung

KI ist neben ihren vielfach hervorgehobenen Potenzialen untrennbar mit erheblichen Risiken verbunden. Diese betreffen nicht nur technisch bedingte und prinzipiell antizipierbare Gefahren, sondern insbesondere unintendierte, in ihrer Tragweite häufig erst im Nachhinein erkennbare sowie mitunter irreversible Folgen (Cheatham et al., 2019; Diaz-Bone & Schrör, 2022; Hirsch-Kreinsen & Krokowski, 2024a; Mayer et al., 2020; Taeihagh, 2021; Veale & Brass, 2019; Zhang et al., 2022). Diese inhärente Unsicherheit bezüglich der Chancen und Risiken lässt einen umfassenden Technikoptimismus (übrigens ebenso wie einen vorschnellen Technikpessimismus) problematisch erscheinen, der zu einer voreiligen und flächendeckenden KI-Einführung führen kann, während potenzielle negative Konsequenzen unzureichend reflektiert werden (Danaher et al., 2017; Elish & boyd, 2018; Nordström, 2022). Häufig basiert die Rechtfertigung des Technologieeinsatzes auf einer Kosten-Nutzen-Abwägung. Dabei tragen jedoch nicht selten andere Personengruppen als die Nutznießenden die negativen Folgen, was auf eine strukturelle Asymmetrie verweist (Crawford, 2021; Kalluri, 2020; Stilgoe, 2020; Veale et al., 2023; Weinberg, 2022). So profitierte etwa im dargestellten Parcoursup-Beispiel die staatliche Bildungsverwaltung von einem effizienteren und administrativ entlastenden Zulassungsverfahren, während die Studienbewerber:innen unter den diskriminierenden Konsequenzen litten (Carpentier & Courtois, 2025; Frouillou et al., 2020). Weitere Ungleichgewichte zeigen sich etwa zwischen Bildungstechnologieunternehmen, die ökonomische Vorteile aus der KI-

basierten Verarbeitung personenbezogener Daten erzielen, und den anwendenden Bildungsinstitutionen, die für Datenschutz und Sorgfaltspflichten verantwortlich bleiben (Williamson, 2024).

Zahlreiche bereits dargelegte Fallbeispiele diskriminierender KI-Anwendungen verdeutlichen, dass der Einsatz von KI tief in gesellschaftliche Ordnungen eingreifen kann. Solche Fälle führen spätestens mit dem Sichtbarwerden negativer Konsequenzen zu umfassender Kritik und öffentlichen Protesten, angeführt von den Geschädigten, die als Ausdruck nachträglicher Politisierung und mangelnder Legitimation verstanden werden (Gallego & Kurer, 2022; König & Wenzelburger, 2021; Ulbricht, 2020; Wieringa, 2023). Dabei kann Legitimität nicht nur dann infrage gestellt werden, wenn Systeme fehlerhaft arbeiten, sondern auch dann, wenn ihre Eignung als legitimes Mittel zur Erreichung politischer Ziele bestritten wird (Simmons, 2018). So weist König (2018) zudem auf das Spannungsverhältnis zwischen Input- und Output-Legitimität hin: Selbst effiziente und wirksame KI-Systeme können demokratietheoretisch problematisch sein, wenn Betroffene keine Möglichkeit hatten, Einfluss auf deren Einführung zu nehmen oder ihren Einsatz zu kontrollieren und im Zweifel zu widerrufen.

Da KI jedoch zunehmend im öffentlichen Raum eingesetzt wird und ein Entzug des Einflusses der Systeme mancherorts kaum mehr möglich erscheint, handelt es sich bei der Einführung von KI nicht um eine rein technische Frage, sondern um eine in besonderem Maße rechtfertigungspflichtige Policy-Entscheidung (Christen et al., 2020; Züger & Asghari, 2023). Entsprechend sind Fragen der KI-Legitimation und -Regulierung von doppelter Relevanz: erstens gesellschaftspolitisch, angesichts der ungleichen Verteilung von Kosten und Nutzen sowie der Möglichkeit irreversibler Folgen (Gallego & Kurer, 2022; Kalluri, 2020; Mensch & Schmidt, 2003; Stilgoe, 2020), und zweitens demokratietheoretisch, aufgrund der Gefahr eines Legitimitätsverlustes, der nicht nur die Technologie, sondern auch politische Entscheidungsträger:innen, Entwickler:innen und Anwender:innen betreffen kann (Chai et al., 2022; Nullmeier et al., 2012; Schiff et al., 2022; Weyer, 2008).

Die politische Steuerung von KI ist jedoch durch die hohe Unsicherheit und Komplexität ihres Einsatzes erschwert (Danaher et al., 2017; Gasser & Almeida, 2017; Lewallen, 2021; Nordström, 2022; Perry & Uuk, 2019; Taeihagh, 2021). KI unterscheidet sich dabei von anderen Risikotechnologien wie Atomkraft oder Gentechnik durch ihre Vielseitigkeit, die eine Vielzahl politischer Ressorts betrifft und die Policy-Gestaltung entsprechend komplexifiziert (Lemke et al., 2024; Novelli et al., 2024). Hinzu kommt die Unsicherheit über die tatsächlichen Auswirkungen des Technikeinsatzes, deren politische Konsequenzen im bereits erwähnten

Collingridge-Dilemma zum Ausdruck kommen (Collingridge, 1985; Genus & Stirling, 2018; Kudina & Verbeek, 2019). Diese Ausgangslage eröffnet einen Aushandlungsraum, in dem unterschiedliche Akteur:innen konkurrierende Deutungen von Potenzialen und Risiken ins Feld führen, um ihre jeweiligen Technologie- und Regulierungsziele zu legitimieren (Gamerdinger & Willers, 2025; Obendiek & Seidl, 2023; Rahm, 2023b).

Denn KI-Systeme sind weder neutrale Werkzeuge noch bloß technische Artefakte; vielmehr sind sie in ihrer Entwicklung und Nutzung von normativen Annahmen und Wertvorstellungen durchdrungen. Die Definition von Fairness etwa ist keine technische, sondern eine politische Frage, deren Beantwortung moralische Werturteile voraussetzt (Wong, 2020). Ebenso ist die Entscheidung für den Einsatz von KI nicht als rein technische, sondern als genuin politische Frage zu begreifen, die tief in bestehende gesellschaftliche Ordnungen eingreift (Green, 2021a, 2021b; Li, 2024; Totschnig, 2019). Werden die erforderlichen Aushandlungsprozesse zur Klärung dieser Fragen depolitisiert, droht die Entstehung oder Verfestigung technokratischer Machtstrukturen, wodurch die demokratische Legitimität des Technologieeinsatzes unterminiert wird (Bartl, 2023).

Verstärkt wird dieses Risiko durch eine öffentliche Debatte, die vielfach von einem positivistischen, teils solutionistischen Verständnis von KI geprägt ist (Cools et al., 2024; Köstler & Ossewaarde, 2022; Nemorin et al., 2023; Ossewaarde & Gülenç, 2020; Sun et al., 2020). Im Mittelpunkt steht dabei häufig die Frage, welche gesellschaftlichen Probleme durch KI gelöst werden können und wie ihr Einsatz technisch zu gestalten ist. Kritischere Perspektiven beschränken sich hingegen oftmals auf das Erkennen und Korrigieren fehlerhafter Systeme. Da unbeabsichtigte Folgen jedoch meist erst nach der Einführung von KI und dem Eintreten negativer Konsequenzen sichtbar werden (König & Wenzelburger, 2021; Ulbricht, 2020), ist eine vorgelagerte Prüfung erforderlich, ob KI-Anwendungen überhaupt ein legitimes und moralisch vertretbares Mittel zur Bearbeitung gesellschaftlicher Probleme darstellen. Dies gilt insbesondere in Bereichen wie der Bildung, wo sie weitreichende Konsequenzen für individuelle Lebensverläufe sowie für bestehende Macht- und Gerechtigkeitskonstellationen nach sich ziehen können (Diaz-Bone & Schrör, 2022; Katzenbach & Ulbricht, 2019; Morozov, 2013; Smuha, 2022; Teräs et al., 2020; Züger & Asghari, 2023).

Dieser Perspektivwechsel erfordert sowohl eine Refokussierung der öffentlichen Debatte als auch der Forschung, die bislang primär outputorientiert ausgerichtet ist und vor allem Fragen der Legitimitäts-, Akzeptanz- oder Fairnesswahrnehmung KI-generierter Entscheidungen (König & Wenzelburger, 2022; Lünich & Keller, 2024; Marcinkowski et al., 2020; Martin &

Waldman, 2022; Shin, 2021; Starke et al., 2022; Starke & Lünich, 2020) sowie Aspekte der technischen Diskriminierungsfreiheit und Erklärbarkeit selbiger untersucht (Altman et al., 2018; Angelov et al., 2021; Barredo Arrieta et al., 2020; Gunning et al., 2019; A. Holzinger et al., 2019; Mehrabi et al., 2022; Minh et al., 2022). Demgegenüber bleibt die Frage nach der Legitimität der politischen Entscheidung über die Einführung von KI und ihre moralischen Grundlagen weitgehend unberücksichtigt (Tallberg et al., 2023). Forschung und ethische Leitlinien adressieren vornehmlich die Ausgestaltung bestehender Systeme (*wie* KI eingesetzt wird), während die vorgelagerte Grundsatzentscheidung (*ob* KI eingesetzt werden sollte) selten problematisiert wird. Anstelle einer *ex-post-orientierten* Erklärung der KI-Entscheidung ist daher eine *ex-ante-Betrachtung* der Legitimität des KI-Einsatzes erforderlich, um zum einen die ethisch fragwürdige Inkaufnahme potenzieller Risiken zu vermeiden, zum anderen den Anforderungen demokratischer Entscheidungsfindung zu entsprechen und schließlich das Risiko von Vertrauensverlust und öffentlicher Ablehnung zu reduzieren.

Ohne eine solche Betrachtung droht die Debatte, in eine Solutionismus-Falle zu tappen, die Selbst et al. (2019) wie folgt beschreiben: „Failure to recognize the possibility that the best solution to a problem may not involve technology" (S. 63). Dies kann zu einer unhinterfragten Annahme führen, dass technologische Lösungen als selbstverständliche Problemlöser der Gegenwartsgesellschaft angenommen werden (Morozov, 2013). boyd und Crawford (2012) warnen daher schon früh: „Just because it is accessible does not make it ethical" (S. 671).[2]

Vor diesem Hintergrund verfolgt die vorliegende Arbeit einen konstruktivistisch-pragmatischen Ansatz und stellt die forschungsleitende Frage, auf welcher moralischen Grundlage – trotz der erheblichen und teils unabsehbaren Risiken – in der bildungspolitischen Debatte in Deutschland eine Rechtfertigung und damit Legitimation des Einsatzes von KI erfolgt. Diese Frage wird im Rahmen der theoretischen Ausarbeitung weiter präzisiert und an späterer Stelle in drei empirisch zu beantwortende Forschungsfragen überführt (siehe Kapitel 7). Das folgende Kapitel widmet sich der theoretischen Konzeption der Herstellung politischer Legitimität solcher Entscheidungen, bevor daran anschließend die Bedeutung der Politisierung sowie der Rechtfertigung und Kritik für die Legitimation der KI-Einführung dargelegt wird.

[2] Die Gefahr einer technokratisch-solutionistischen Anwendung von KI, verbunden mit strukturellen Machtungleichgewichten, ist mittlerweile Gegenstand einer wachsenden technikkritischen Auseinandersetzung. Auf diese wird ebenso wie auf die definitorische Spezifizierung des Solutionismus-Begriffs im weiteren Verlauf dieser Arbeit (siehe Abschnitte 6.1 und 6.4.4) noch vertieft eingegangen.

3. Theoretische Grundlagen politischer Legitimation im Kontext der Einführung Künstlicher Intelligenz

Die Untersuchung der Legitimation der politischen Entscheidung über die KI-Einführung erfordert eine vertiefte Auseinandersetzung mit den theoretischen Grundlagen und Mechanismen politischer Legitimitätsherstellung. Zu diesem Zweck werden zunächst die der Arbeit zugrunde liegenden Begriffsverständnisse von Legitimität und Legitimation dargelegt. Anschließend werden mit der Unterscheidung normativ-theoretischer und empirisch-analytischer Strömungen zwei unterschiedliche konzeptionelle Forschungsperspektiven auf politische Legitimität skizziert. Diese Differenzierung bildet die Grundlage, um darauf aufbauend ein drittes, verbindendes pragmatisches Legitimitätsverständnis darzulegen, das sowohl die moralisch-normative Fundierung politischer Legitimation berücksichtigt als auch die empirische Rekonstruktion der Legitimitätspraktiken politischer Akteur:innen ermöglicht. Das pragmatische Verständnis, auf das sich diese Arbeit stützt, erlaubt es somit, die Legitimation der politischen Entscheidung über den Einsatz von KI konzeptuell als Aushandlung politischer, moralisch fundierter Rechtfertigungen zu fassen und die zugrunde liegenden Mechanismen in der späteren empirischen Analyse systematisch zu rekonstruieren.

Dabei zählt Legitimität zu den meistdiskutierten Konzepten der politikwissenschaftlichen Theorie und Forschung (zur begriffshistorischen Entwicklung siehe Glaser, 2013). Im Kern geht es um die Frage nach der Anerkennungswürdigkeit politischer Herrschaft und Ordnung (Beetham, 1991/2013; Blatter, 2007; Habermas, 1976a). Legitimität fungiert als Orientierungsrahmen für soziales Handeln, indem sie politische Herrschaft in einen gesellschaftlich akzeptierten Ordnungsrahmen einbettet (Diehl, 2015; Habermas, 1976a; Luckmann, 2001). Damit erklärt sich ihre grundlegende Bedeutung für die Funktionsfähigkeit demokratischer Systeme: Indem sie die Anerkennung politischer Ordnung gewährleistet, sichert sie die Systemstabilität sowie die Bereitschaft des Volkes zur Befolgung der Ordnung (Nullmeier et al., 2012; Scharpf, 2004). Beetham (1991/2013) fasst diese demokratische Bedeutung zusammen:

> Why does legitimacy matter? It matters because of the difference it makes to people's attitudes and behaviour. To the extent that people acknowledge power as rightful, as validly acquired and properly exercised, they will feel a corresponding obligation to obey and support it without having to be bribed or coerced into doing so. (S. xi)

Wie Beetham hervorhebt, geht der Legitimitätsanspruch über eine normative Erwartung hinaus und entfaltet verhaltenswirksame Kraft. Dies zeigt sich nicht nur in Bezug auf die lange Zeit im Fokus der Legitimitätsforschung stehende Anerkennung politischer Strukturen (*Polity*) und Prozesse (*Politics*), sondern bereits auf Ebene einzelner Policy-Entscheidungen. So kann sich

mangelnde Legitimität im Widerstand gegen einzelne politische Maßnahmen und in der Folge im Scheitern selbiger widerspiegeln (Wallner, 2008). Die Ablehnung spezifischer politischer Entscheidungen wird bereits als früher Indikator eines beginnenden Legitimitätsverlusts des politischen Systems betrachtet, was zugleich ein zunehmendes Interesse an der Untersuchung der Policy-Legitimität begründet (Nullmeier et al., 2012), wie sie im Mittelpunkt der vorliegenden Arbeit steht. Die fehlende Akzeptanz politischer Maßnahmen verweist oftmals auf tieferliegende Spannungen zwischen der gesellschaftlichen Anerkennung normativer Ansprüche und ihrer praktischen Umsetzung, wie sie bereits im Kontext der KI-Einführung sichtbar geworden sind. Die bereits angeführten Beispiele diskriminierender KI-Anwendungen verdeutlichen, dass politische Legitimität nicht allein durch den normativen Anspruch auf Fairness oder Transparenz gesichert ist, sondern davon abhängt, ob diese Prinzipien in der praktischen Umsetzung des KI-Einsatzes tatsächlich eingelöst werden. Wenn KI-Systeme trotz gegenteiliger Versprechen diskriminierende Effekte erzeugen oder gesellschaftliche Ungleichheiten verstärken, gerät die politische Legitimation ihrer Einführung ins Wanken, was sich im Widerstand gegen KI-basierte Politiken und in der Folge im Scheitern ihrer Umsetzung widerspiegeln kann (Danaher, 2016; Gutiérrez, 2023; Ulbricht, 2020; Wieringa, 2023).

3.1. Zur Abgrenzung statischer Legitimität und dynamischer Legitimation

Zugleich verdeutlichen diese Beispiele, dass die Legitimität der politischen Entscheidung über den Einsatz von KI nicht dauerhaft gesichert ist, sondern „stets aufs Neue erzeugt, kommuniziert und erhalten werden“ (Nullmeier, 2010, S. 9) muss. Legitimität wird somit an dieser Stelle nicht als historisch gewachsener und institutionell verfestigter Zustand begriffen, sondern als dynamischer, kontinuierlich zu reproduzierender Prozess. Analytisch lässt sich die Abgrenzung zwischen einem statischen und einem prozessualen Legitimitätsverständnis produktiver fassen, wenn zwischen Legitimität und Legitimation unterschieden wird. Während Legitimität einen normativen Zustand politischer Anerkennungswürdigkeit beschreibt, bezeichnet Legitimation die gesellschaftlichen Prozesse, durch die dieser Zustand hergestellt wird (Habermas, 1976b; Hennis, 1976a).

Dieser Prozesscharakter der Herstellung von Legitimität wird deutlich in der vielzitierten Definition Suchmans (1995): „Legitimacy is a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions“ (S. 574). Diese Definition verdeutlicht somit, dass Legitimität erstens in sozialen Prozessen konstruiert und zweitens normativ fundiert wird. Für die vorliegende Arbeit ist die Analyse dieser Legitimationsprozesse zentral, da sie ermöglicht, die Legitimität der politischen Entscheidung über die KI-Einführung nicht retrospektiv als

dauerhaft gegebenen Zustand vorauszusetzen, sondern prospektiv empirisch zu rekonstruieren, wie Akteur:innen sie durch spezifische Technikdeutungen und Formen der Rechtfertigung herstellen und somit versuchen, Anerkennung derselbigen in der Öffentlichkeit zu erzeugen (Beckman et al., 2024; Conti & Seele, 2023; Geels & Verhees, 2011; Korneeva et al., 2023). Um diese Legitimationsbestrebungen analytisch zu fassen, ist es notwendig, im Folgenden die Mechanismen der Legitimitätsherstellung näher zu beleuchten.

3.2. Forschungsperspektiven auf die Mechanismen der Legitimitätsherstellung

In der Forschungsliteratur differieren die Perspektiven darauf, wie politische Legitimität erzeugt wird. Dabei zeigt sich ein Spannungsverhältnis zwischen (a) normativ-theoretischen und (b) empirisch-analytischen Strömungen, wobei in jüngerer Zeit (c) pragmatische Ansätze dieses Spannungsfeld zu überbrücken versuchen (Diehl, 2015; Kneip & Merkel, 2020; Sarcinelli, 1998; Wiesner & Harfst, 2019). Mit diesen Perspektiven verknüpft sind verschiedene Deutungsansätze der Mechanismen politischer Legitimation. Darunter befinden sich repräsentative[3], deliberative und verfahrensbezogene Ansätze (Diehl, 2015). Die Strömungen und Mechanismen werden nachfolgend kurz umrissen, um auf dieser Basis die Notwendigkeit eines pragmatischen Legitimitätszugangs zu begründen, der es erlaubt, die politische Legitimation des KI-Einsatzes als dynamischen Prozess öffentlicher Rechtfertigung zu verstehen.

Aus der Perspektive der Vertreter:innen einer politikphilosophischen *normativ-theoretischen Strömung* (u. a. Forst, 2014, 2015; Habermas, 1976b; Rawls, 1996) ergibt sich Legitimität aus einem normativ gerechtfertigten Geltungsanspruch politischer Herrschaft, der sich auf demokratische Prinzipien wie Volkssouveränität, Gewaltenteilung und partizipative Willensbildung stützt (Sarcinelli, 1998). Diese Sichtweise erhebt einen hohen Anspruch an die normative Qualität politischer Entscheidungsprozesse, wie etwa im deliberativen Legitimitätsverständnis sichtbar wird, welches insbesondere durch Habermas (1976a), Cohen (2003) und Rawls (1971, 1996) vertreten wird. In diesem kommunikativen Ansatz gelten politische Entscheidungen als legitim, wenn sie im Rahmen offener, herrschaftsfreier Deliberation zustande kommen. Legitimität erfordert dabei den Einbezug aller relevanten Positionen in einen rationalen Abwägungsprozess, wobei gemäß dem „zwanglosen Zwang des besseren Argumentes“ (Habermas, 1992, S. 367) ein allseits anerkannter Konsens erzielt werden soll. Auf der einen Seite hebt dieses Ideal die Bedeutung der kommunikativen Aushandlung der Legitimität politischer Entscheidungen hervor und ist damit für Fragen der Legitimation des KI-Einsatzes normativ bedeutsam.

[3] Repräsentative Modelle leiten Legitimität aus der gesellschaftlichen Delegation politischer Entscheidungsgewalt auf gewählte Repräsentant:innen ab (Diehl, 2015; Pitkin, 1967). Für die vorliegende Analyse ist dieser Ansatz jedoch nur begrenzt anschlussfähig, da er Legitimität politischer Entscheidungen im Wesentlichen auf demokratische Autorisierung reduziert und damit die soziale Konstruktion der Policy-Entscheidung ausblendet.

Auf der anderen Seite erweist sich das Ideal eines herrschaftsfreien Diskurses und rationalen Konsenses jedoch in der Praxis als empirisch kaum einlösbar, da reale politische Diskurse unter anderem von Machtasymmetrien, differierenden Artikulationsfähigkeiten und konfliktiven Aushandlungen geprägt sind (Bächtiger & Wyss, 2013; J. Johnson, 1998; Mouffe, 2015/2005; Shapiro, 1999). Der deliberative Ansatz beschreibt somit eine idealtypische, aber selten erreichte Form politischer Entscheidungsfindung und eignet sich daher nur eingeschränkt zur Analyse politischer Legitimationspraktiken, weshalb vermehrt empirische Legitimationszugänge gefordert werden (Gadinger & Yildiz, 2012).

So verweisen im Gegensatz dazu Vertreter:innen der sozialwissenschaftlichen *empirisch-analytischen Strömung* auf den legitimatorischen Gehalt gesellschaftlicher Anerkennung. In Anlehnung an Weber (2011/1956) wird sie als „Legitimitätsglaube" verstanden, das heißt als sozial wirksamer Glaube an die Rechtmäßigkeit politischer Ordnung und Entscheidungen, der sich durch Meinungen, Einstellungen und Wertüberzeugungen empirisch erfassen lässt (Kneip & Merkel, 2020; Sarcinelli, 1998; Schmelzle, 2012). Legitimität gilt in diesem Verständnis als gegeben, wenn sie von den Adressat:innen als solche wahrgenommen wird (Kneip & Merkel, 2020; Suchman, 1995). Damit rücken Fragen der Operationalisierung des Legitimationsprozesses und somit vor allem verfahrensbasierte Legitimitätsverständnisse in den Forschungsfokus, die an den Legitimitätsglauben anknüpfen. In diesem Kontext betont Luhmann (1969/1983), dass Legitimität als Resultat institutionalisierter Abläufe und der routinisierten Anerkennung formaler Entscheidungen durch etablierte Verfahren zu verstehen ist. So beruhe Legitimität „auf einem sozialen Klima, das die Anerkennung verbindlicher Entscheidungen als Selbstverständlichkeit (…) der Geltung der amtlichen Entscheidung ansieht" (S. 34). Die politische Entscheidung des KI-Einsatzes gilt demnach als legitim, wenn sie durch etablierte und transparente Verfahren zustande kommt und aufgrund dessen gesellschaftlich als selbstverständlich akzeptiert wird.

Analytisch umfasst Luhmann damit Faktoren, die als Input-Legitimität bezeichnet werden können, welche unter anderem von Easton (1965, 1975) und Scharpf (1998, 1999) durch die Dimension der Output-Legitimität ergänzt wird. Input-Legitimität bezeichnet dabei die demokratische Anerkennung politischer Entscheidungen, die aus der Partizipation der Bürger:innen im politischen Willensbildungsprozess hervorgeht: „Political choices are legitimate if and because they reflect the ‚will of the people' – that is, if they can be derived from the authentic preferences of the members of a community" (Scharpf, 1999, S. 6). Die Output-Legitimität misst sich hingegen an der Effektivität politischer Problemlösungen und am Nutzen für das Gemeinwohl (Blatter, 2007; Easton, 1975; Glaser, 2013; Scharpf, 1999). In der

Weiterentwicklung dieses Konzepts wird ergänzend auf die Bedeutung der Throughput-Legitimierung verwiesen – darunter auch im Kontext KI-basierter Entscheidungsfindung (Starke & Lünich, 2020) –, welche die Rolle der politischen Prozesse in der Verarbeitung des Inputs zum eigentlichen Policy-Output hervorhebt (Schmidt, 2013; Schmidt & Wood, 2019). Mit dieser differenzierten Betrachtungsweise eröffnet sich ein Spannungsfeld zwischen einer normativ-demokratischen Input-Legitimität und einer mitunter technokratisch effizienzorientierten Output-Legitimität (Caramani, 2020; Scharpf, 1999), das insbesondere im Kontext der politischen Entscheidung über den KI-Einsatz virulent wird.

So unterliegt die Entscheidung über den Einsatz von KI-Systemen sowohl normativen Anforderungen, etwa an demokratische Beteiligung (Input), als auch an die erwartete Wirksamkeit und Gemeinwohlorientierung (Output). Besonders im Hinblick auf die Input-Dimension betonen einige Autor:innen die Notwendigkeit partizipativer Verfahren, um die gesellschaftliche Legitimation solcher Entscheidungen sicherzustellen (z. B. Danaher et al., 2017; Lee et al., 2019). Die Gestaltung und KI-Einführung werden demnach zunehmend als Aufgabe kollektiver Aushandlung verstanden, die nicht allein durch technische Expertise legitimiert werden kann. Gleichzeitig zeigt sich auf der Output-Seite ein wachsender Legitimationsdruck, den Einsatz von KI-Technologien nicht nur technisch oder ökonomisch, sondern insbesondere im Hinblick auf ihren Nutzen für das Gemeinwohl zu rechtfertigen. In diesem Kontext ist die Forderung von Züger und Asghari (2023) zu verorten, dass sich legitime KI-Entscheidungen an einem gemeinwohlorientierten Leitbild orientieren und ihren Beitrag zur Förderung gesellschaftlicher Werte offenlegen müssen. Derartige Forderungen verweisen darauf, dass politische Legitimität zunehmend an die explizite Ausrichtung politischer Maßnahmen an normativen Zielvorstellungen gebunden wird. Im Kontext technologischer Innovation treten die Logiken von Input und Output jedoch häufig in Konflikt (König, 2018; Sætra, 2020), etwa wenn demokratisch nicht legitimierte Expert:innengremien KI-Anwendungen mit weitreichenden gesellschaftlichen Folgen entwickeln. Die Einführung von KI wirft somit nicht nur technische, sondern ebenso politische Fragen auf, wenn etwa danach gefragt wird, auf welcher normativen Grundlage bestimmte Anwendungen erlaubt oder verboten werden.

Vor diesem Hintergrund wird die Notwendigkeit der Integration normativer und empirischer Perspektiven in einem pragmatischen Ansatz deutlich. Dabei wird gefordert, die empirische Legitimitätsforschung auf eine theoretisch-normative Basis zu stellen (Harfst & Wiesner, 2019; Kneip & Merkel, 2020). Während normative Ansätze Erwartungen an rechtmäßige Herrschaft und Diskurse formulieren und empirische Ansätze in den Blick nehmen, ob diese Erwartungen in der Gesellschaft als legitim anerkannt werden, bleibt in beiden Fällen der kommunikative

Legitimationsprozess weitgehend unberücksichtigt. Komplexe Entscheidungen über den KI-Einsatz lassen sich weder durch institutionelle Verfahren noch durch demokratische Delegation oder deliberative Ideale hinreichend erfassen. Für die Analyse der Legitimation der politischen Entscheidung über den KI-Einsatz bedarf es daher eines Ansatzes, der normative und empirische Dimensionen verbindet und Legitimation als kommunikativen, sozial konstruierten Prozess begreift. Ein solcher pragmatisch-kommunikativer Zugang versteht Legitimation somit nicht als Zustand, sondern als Ergebnis eines dynamischen Aushandlungsprozesses öffentlicher Rechtfertigung. Er bildet damit den theoretischen Ausgangspunkt für den im folgenden Abschnitt dargestellten Ansatz der Soziologie der Konventionen, der diesen kommunikativen Charakter politischer Legitimität systematisch theoretisiert und empirisch anschlussfähig macht.

3.3. Das pragmatische Legitimationsverständnis: Grundlagen der Soziologie der Konventionen

Die Soziologie der Konventionen (auch Soziologie der Kritik, kurz: SdK)[4] bietet als erste systematisch vorgelegte soziologische Analyse alltagsweltlicher Rechtfertigungspraktiken einen fruchtbaren Ansatz für die Untersuchung der Legitimation politischer Entscheidungen (Susen, 2017). Sie versteht Legitimität nicht als abstrakten Geltungsanspruch, sondern als Resultat pluraler, kontextgebundener Aushandlungsprozesse zwischen Akteur:innen, die auf verschiedene moralische Ordnungen zurückgreifen. Damit eröffnet sie einen empirisch fundierten Zugang zur Legitimität politischer Entscheidungen und verbindet normative Grundlagen mit empirisch-analytischen Ansprüchen. Das pragmatische Legitimitätsverständnis der SdK basiert auf einem kommunikativen Grundgedanken, verzichtet jedoch auf idealisierte Voraussetzungen rationaler Deliberation und entwickelt stattdessen einen empirisch-pragmatischen Zugang, der sich gegen den Vorwurf normativer Realitätsferne wendet (Niemann, 2019). In ihrer grundlegenden Konzeption, erstmals ausgearbeitet von Boltanski und Thévenot in *„De la Justification: Les Économies de la Grandeur“* (1991), rückt die SdK alltägliche Praktiken der Rechtfertigung und Kritik ins Zentrum ihrer Analyse. Legitimität wird hier nicht als Ergebnis eines herrschaftsfreien, rationalen Diskurses verstanden, sondern als Resultat eines konflikthaften Aushandlungsprozesses, in dem moralisch begründete Kritik auf gerechtfertigte Ordnungsansprüche trifft (Boltanski & Thévenot, 1991/2018). Einigung wird dabei nicht als vollständiger Konsens konzipiert, sondern als temporärer Kompromiss, der breite Zustimmung erfährt oder zumindest

[4] Die Soziologie der Konventionen (auch *Économie des conventions*) entspringt dem *pragmatic turn* der französischen Sozialwissenschaften der 1980er Jahre und bildet „eine pragmatische Grundlagentheorie für die transdisziplinäre Analyse – nicht nur, aber wesentlich – ökonomischer Institutionen. Zugleich ist sie ein anwendungsbezogener Ansatz, der die empirische Analyse der ökonomischen Koordination und der darin erfolgenden Konstruktionen von Wertigkeiten ermöglicht“ (Diaz-Bone, 2018, S. 2). Für eine detaillierte Einführung in deutscher Sprache siehe ebenda.

von den beteiligten Akteur:innen als hinreichend tragfähig akzeptiert wird. Legitimität entsteht in diesem Verständnis aus situativ begründeten Rechtfertigungen, deren Geltung im Wechselspiel mit kritischen Infragestellungen steht (Bogusz, 2010; Boltanski & Thévenot, 1991/2018; Diaz-Bone, 2018; Gadinger & Yildiz, 2012). Diese Praxis erscheint somit „als ein konstitutives Wechselspiel zwischen Kritik und Bestätigung" (Gadinger & Yildiz, 2012, S. 306), in dem normative Geltungsansprüche empirisch sichtbar und überprüfbar werden.

Im Unterschied zu staatlich zentrierten oder strukturalistisch geprägten Legitimitätsmodellen, die Legitimation als Übereinstimmung mit institutionellen Normen oder funktionalen Systemlogiken begreifen, versteht die SdK Legitimität als nicht-linearen, kontingenten und situativ geprägten Prozess (Smith Ochoa et al., 2021). Ihr Fokus liegt auf der Frage, wie Akteur:innen unter Bedingungen der gesellschaftlichen Komplexität und Unsicherheit soziale Ordnungen (re-)produzieren (Boltanski & Thévenot, 1991/2018). Die SdK versteht Rechtfertigung dabei nicht nur als politisches Phänomen im engeren Sinn, sondern als grundlegende soziale Praxis, die über disziplinäre Grenzen hinweg zur Analyse legitimer Ordnungsbildung beiträgt.

Aus der Kritik an dominanten soziologischen und ökonomischen Theorien (insbesondere an Pierre Bourdieus Theorie habitueller Praxis und an Modellen rationaler Nutzenmaximierung) entwickelt die SdK ein alternierendes Akteursbild: Akteur:innen werden nicht als habituelle Träger:innen von Interessen gefasst, sondern als moralisch urteilende, handlungsfähige Subjekte, die mit normativen Konflikten umgehen und auf Einigung zielen (Boltanski & Thévenot, 1991/2018; Hartmann, 2008). Legitimität wird in diesem Rahmen, anders als noch im Luhmann'schen Verständnis, nicht durch objektivistische Strukturzwänge garantiert, sondern durch subjektiv geteilte normative Überzeugungen getragen (Gadinger & Yildiz, 2012; Hartmann, 2008). Entscheidend ist dabei die Fähigkeit der Akteur:innen, in konkreten Situationen die Angemessenheit von Rechtfertigungen zu beurteilen, Kritik zu artikulieren und moralisch fundierte Kompromisse auszuhandeln, die sich an einer Ordnung des Allgemeinen orientieren (Blättler, 2008; Boltanski & Chiapello, 2003; Gadinger & Yildiz, 2012; Hartmann, 2008). Legitimität beruht demnach auf einer Rangordnung der Allgemeinheit: „Am bedeutendsten ist, was am allgemeinsten ist" (Boltanski & Thévenot, 1991/2018, S. 112). Legitimität kann demnach nicht aus partikularen Interessen abgeleitet werden, sondern bedarf der Bezugnahme auf allgemein anerkannte Werteordnungen, die in einer Situation zur Geltung gebracht werden. Handeln wird somit nicht allein durch instrumentelle Rationalität geleitet, sondern durch normative Prinzipien strukturiert und die Rationalität in konkreten Situationen „durch Konventionen ergänzt und allererst mobilisiert" (Diaz-Bone, 2022, S. 476). In diesem Sinne gelten Akteur:innen als „pragmatisch kompetent" (ebd.).

Die Entscheidung über den KI-Einsatz wird nicht deshalb als legitim wahrgenommen, weil sie formalen Regeln entspricht oder durch ein bestimmtes Verfahren zustande kommt, sondern weil sie sich auf Rechtfertigungen stützt, die als allgemein und moralisch verbindlich gelten. Diese Perspektive erweist sich somit insbesondere im Feld technologiepolitischer Entscheidungsprozesse, wie der Einführung von KI-Systemen, als analytisch besonders anschlussfähig. Denn sowohl die Entwicklung als auch die politische Regulierung solcher Technologien beruhen nicht ausschließlich auf technisch-rationalen Kalkülen, sondern sind in hohem Maße durch normative Vorstellungen, ethische Spannungen und das Ringen um gesellschaftliche Ordnung geprägt (Green, 2021a, 2021b; Li, 2024; Totschnig, 2019). Der theoretische Rahmen der SdK eignet sich daher in besonderer Weise zur Analyse legitimatorischer Dynamiken in der Technologiepolitik: Erstens ermöglicht er es, die normativen Anforderungen an gemeinwohlorientierte Governance, wie etwa im Sinne eines menschenzentrierten Technikverständnisses, empirisch zu rekonstruieren (Tallberg et al., 2023; Züger & Asghari, 2023). Zweitens erlaubt er, eine Analyse politisch-normativer Konflikte, wie sie etwa Mittelstadt (2019) im Bereich der KI-Ethik identifiziert, nicht nur als ethische Dilemmata zu beschreiben, sondern als real ausgetragene Aushandlungsprozesse zu deuten. So kann die Analyse über abstrakte Prinzipien oder die Evaluation von Leitlinien hinausgehen und die konkreten Muster legitimer Entscheidungsfindung sichtbar machen. Und drittens betont die SdK, dass Akteur:innen bestrebt sind, ihre Standpunkte moralisch zu rechtfertigen, wechselseitig abzuwägen und in Bezug auf allgemein anerkannte normative Prinzipien Kompromisse zu schließen. Legitimität beschränkt sich somit nicht auf Konsens oder formale Verfahren, sondern umfasst einen strukturierten Prozess der gesellschaftlich anerkannten moralischen Rechtfertigung, der eine legitime Konfliktlösung ermöglicht.

Bevor die Rolle der politischen Rechtfertigung und ihrer moralischen Grundlage konzeptuell entfaltet wird, ist zu klären, unter welchen Bedingungen solche Rechtfertigungen überhaupt entstehen und wirksam werden. Legitimation setzt voraus, dass politische Entscheidungen öffentlich verhandelt und kritisiert werden können, was erst durch Prozesse der Politisierung möglich wird, in denen ein Thema als strittig und gesellschaftlich relevant markiert wird. Das nächste Kapitel widmet sich daher der Politisierung als Voraussetzung legitimatorischer Aushandlungen, bevor anschließend die konzeptionelle Rolle von Rechtfertigung und ihrer moralischen Grundlage als Mechanismen dieser Legitimitätserzeugung dargestellt wird.

4. Politisierung als notwendige Bedingung legitimer Entscheidungen

Ein kommunikatives Legitimationsverständnis, wie es die SdK vertritt, hebt die zentrale Rolle gesellschaftlicher und politischer Aushandlungsprozesse für eine legitime Entscheidungsfindung hervor. Insbesondere komplexe und potenziell konfliktträchtige politische Entscheidungen, wie jene über die KI-Einführung, erfordern dabei ein gewisses Maß an Politisierung. So ist die Technologiepolitik einerseits durch ein hohes Maß an Unsicherheit und Komplexität gekennzeichnet, während die politischen Entscheidungen über die Entwicklung und den Einsatz von Technik andererseits weitreichende und teils, so in deterministischen Argumentationen angenommen, unumkehrbare Einflüsse auf die Gesellschaft ausüben können (R. Boyd & Holton, 2018; Mensch & Schmidt, 2003; Obendiek & Seidl, 2023). Technik ist dabei in besonderem Maße von Ambivalenz gekennzeichnet, da sie neben beabsichtigten Wirkungen immer auch unerwünschte, unintendierte Folgen mit sich bringen kann, sodass der Umgang mit Technik stets einem Risiko ausgesetzt ist (z. B. Beck, 2014; Habermas, 1985; Luhmann, 1992). In Anbetracht ihrer hohen Wirkungskraft auch auf zukünftige Generationen sowie ihrer (in Teilen) Irreversibilität unterliegt die Technologiepolitik einer besonderen Rechtfertigungspflicht, um die Legitimität des Technikeinsatzes zu gewährleisten: „Wenn schon weitreichende Entscheidungen politisch gefällt werden müssen, ist der Anspruch an deren Verantwortbarkeit und Legitimität zu Recht hoch“ (Mensch & Schmidt, 2003, S. 14). Die Technologiepolitik steht somit vor der Herausforderung einer Gratwanderung zwischen Wirkungsmaximierung der Technologie auf der einen Seite und Risikominimierung auf der anderen Seite. So sollen die der Technologie zugesprochenen Potenziale möglichst umfänglich genutzt werden, während die Risiken eines technischen Fehlschlags mit erheblichen sozialen, ökologischen, ökonomischen und politischen Folgen minimiert werden sollen. In diesem Ansinnen bewegt sich die Technologiepolitik stets im Spannungsfeld zwischen Innovationspolitik, die auf die wirtschaftliche Förderung technischer Neuerungen abzielt, und Technikregulierung, die deren (unintendierte) Risiken und gesellschaftliche Auswirkungen kontrollieren soll (Fisher et al., 2006; Mensch & Schmidt, 2003; Weyer, 2008; Willems, 2021).

Vor diesem Hintergrund argumentieren Lemke et al. (2024), dass eine robuste Politisierung eine notwendige Bedingung effektiver Entscheidungsfindung im Umgang mit komplexen Problemlagen darstellt.[5] Die Autor:innen verweisen auf die von Lewallen (2021) beschriebene

[5] Die Autor:innen fassen „robuste Politisierung“ als effektive Entscheidungsfindung unter Deutungshoheit politischer Entscheider:innen auf. Dieses der Public-Governance-Literatur entstammende Verständnis betont die Steuerungsfähigkeit des Staates (vgl. Capano & Woo, 2017; Sørensen & Ansell, 2023), jedoch zulasten eines breiten, kontroversen Diskurses. Eine solche Perspektive läuft Gefahr, Politisierung primär als Mittel politischer Stabilisierung zu begreifen. Gleichwohl kann der Annahme der Autor:innen gefolgt werden, dass Politisierung eine zentrale Bedingung legitimer Entscheidungen im Kontext komplexer Technologien darstellt.

Problemdefinitions-Unsicherheit in frühen Phasen der Regulierung neuer Technologien: Etablierte Problembeschreibungen werden destabilisiert, während neue Deutungsmuster noch nicht gefestigt sind. Die schnelle Verbreitung von KI-Anwendungen erzeugt ein Moment der Kontingenz, der neue regulatorische Spielräume eröffnet und die Notwendigkeit öffentlicher Problemdefinitionen verstärkt. In solchen Phasen steigt die Bedeutung öffentlicher Debattenräume, in denen konkurrierende Deutungen ausgehandelt und legitimiert werden (Schmidt, 2002a). Dabei versuchen sowohl politische Entscheidungsträger:innen als auch zivilgesellschaftliche Akteur:innen, ihre je eigene Problemdefinition im Rahmen des Politisierungsprozesses durchzusetzen. Dieser Prozess bildet maßgeblich die Grundlage für die Legitimation künftiger Regulierungsmaßnahmen (Lemke et al., 2024; Sørensen & Ansell, 2023).

Diese Überlegungen zur Rolle der Politisierung verweisen auf eine tieferliegende Verbindung zwischen politischen Entscheidungsprozessen und gesellschaftlichen Rechtfertigungskonflikten. Denn die Frage, wie Legitimität in technologischen Transformationsprozessen hergestellt wird, lässt sich nicht allein aus der Perspektive staatlicher Steuerung beantworten. Vielmehr verweist ein kommunikativer Legitimationsbegriff, wie er etwa in der SdK vertreten wird, auf die zentrale Rolle öffentlicher Aushandlungsprozesse für die Herstellung politischer Geltung. So sehen sich politische Institutionen in ihrer Rolle der Wirklichkeitskonstruktion und Komplexitätsreduktion einem ständigen Wechselspiel zwischen Kritik und Rechtfertigung ausgesetzt. Demnach müssen sie immer wieder auf die ihnen entgegengebrachten Kritiken reagieren und ihren Standpunkt verdeutlichen beziehungsweise bestätigen (Boltanski, 2008). In diesem Verständnis ist Politisierung nicht nur Voraussetzung effektiver Entscheidungsvorbereitung, sondern ebenso konstitutiver Bestandteil legitimer politischer Praxis.

4.1. Politisierung als kontingenter politischer Aushandlungsprozess

Vor diesem Hintergrund ist es geboten, den Begriff der Politisierung theoretisch zu präzisieren. Eine Minimaldefinition schlägt hierzu Zürn (2013) vor, demzufolge Politisierung als „die Forderung nach oder der Akt des Transports einer Entscheidung oder einer Institution in den Bereich des Politischen“ (S. 13) verstanden werden kann. Diese Definition erlaubt zwei zentrale Schlüsse: Erstens gilt, dass alle Themen, die im politischen Raum verhandelt werden, bereits politisiert sind. In Bezug auf die KI-Debatte in Deutschland könnte somit KI etwa seit der ersten Thematisierung im Deutschen Bundestag im Jahr 1984 als politisiert gelten (Willems, 2021). Zweitens impliziert diese Perspektive jedoch auch eine differenzierungstheoretische Herangehensweise, wie sie Zürn (2013) entlang zweier zentraler gesellschaftlicher Achsen entfaltet. So kann die Beschreibung des politischen Raums, analog zur Legitimitätsdebatte, sowohl aus einer systemtheoretischen Beschreibung – allen voran nach Luhmann (1990) – als auch aus einer

diskurstheoretischen Differenzierungstheorie des politischen Systems erfolgen. Folgt man jedoch zunächst den Gedanken Luhmanns, so muss das politische System als eines von vielen gesellschaftlichen Systemen begriffen werden, welches einer ihm inhärenten Funktionslogik folgt und das Ziel von kollektiv bindenden Entscheidungen verfolgt (Luhmann, 1986). Im Zuge von Politisierungsprozessen werden dann Sachverhalte anderer Teilsysteme, wie beispielsweise der Wirtschaft oder der Wissenschaft, die wiederum eigenen Funktionslogiken folgen, in das politische System transportiert. Diese Annahme sieht sich jedoch der Kritik ausgesetzt, gesamtgesellschaftliche Zusammenhänge unzureichend einzubeziehen, sodass Politisierung „in dieser systemtheoretischen Sichtweise primär ein Problem der sektoralen Zuweisung von Zuständigkeiten" (Zürn, 2013, S. 15) bleibt.

Die Abbildung gesamtgesellschaftlicher Zusammenhänge ist ein zentrales Anliegen des diskurstheoretischen Standpunkts, der neben der funktionalen Betrachtung des politischen Systems, kollektiv bindende Entscheidungen zu treffen, darüber hinaus der Aushandlung von Werten und Normen beziehungsweise der Definition des normativ Wünschenswerten einen hohen Stellenwert zuschreibt (vgl. Abschnitt 3.2). Im Zentrum dieser Überlegungen steht, anknüpfend an das pragmatische Legitimationsverständnis, die Aushandlung von Kompromissen, in denen beteiligte Akteur:innen ihre Standpunkte rechtfertigen und kritisieren (Boltanski & Thévenot, 1991/2018). Die so entstehende Öffentlichkeit setzt sich, wie bereits angedeutet, aus verschiedensten Akteursgruppen zusammen, wie Vertreter:innen der Zivilgesellschaft, der Wissenschaft, der Medien und der Politik (Habermas, 1992). Zugleich bilden diese öffentlichen Aushandlungsprozesse die Basis der Entstehung des politischen Raums. So kann von Politisierungsprozessen gesprochen werden, wenn einzelne Teilsysteme eigene Anliegen in die Öffentlichkeit tragen, sodass im politischen Raum reflektiert werden kann, „in welchem gesellschaftlichen Teilsystem und nach welcher Funktionslogik bestimmte Entscheidungen gefällt werden sollen" (Zürn, 2013, S. 17). Diesem Verständnis liegt demnach nicht ausschließlich die bloße (systemtheoretische) Verschiebung von Sachverhalten in den politischen Raum zugrunde, sondern vielmehr die Reflexion der Angemessenheit von Entscheidungsbefugnissen oder Entscheidungen selbst.

Vor dem Hintergrund dieser beiden Differenzierungstheorien erweitert Zürn (2013) die zu Beginn des Kapitels angeführte Definition der Politisierung um ein verbindendes Element beider Perspektiven. Dabei schließt er explizit alle Akteur:innen ein, die entweder bereits Teil der Öffentlichkeit sind, wie beispielsweise Politiker:innen, Expert:innen oder Interessenvertreter:innen, oder aber auch zum Teil der Öffentlichkeit werden können, indem sie fähig sind, einen öffentlichen Protest zu organisieren, beispielsweise, weil sie von KI-Entscheidungen

unmittelbar betroffen sind. Als Ort der Politisierung werden ferner sowohl das systemtheoretische als auch das diskurstheoretische Verständnis des Politischen erklärt:

> Politisierung soll definiert werden als der Prozess, mittels dessen Entscheidungskompetenzen und die damit verbundenen autoritativen Interpretationen von Sachverhalten in die politische Sphäre gebracht werden, d.h. entweder in das politische Teilsystem (definiert durch die politische Funktionslogik) oder in den politischen Raum (definiert durch Debatten über die angemessene Funktionslogik für eine gegebene Problemlage) transportiert werden. (Zürn, 2013, S. 19)

Dieses Verständnis von Politisierung umfasst somit mehr als den bloßen Transport eines Themas in den politischen Raum. Das diskursive Element der Debatte schließt den Moment der normativen Kontroverse über die Legitimität von Entscheidungsbefugnissen und -inhalten mit ein (Forst, 2015; Hay, 2007; Niemann, 2019; Warren, 1999). Politisierung tritt dort auf, wo Entscheidungen selbst zur Disposition gestellt werden, etwa wenn ihre Akzeptanz öffentlich infrage gestellt wird (König & Wenzelburger, 2021; Niemann, 2019; Ulbricht, 2020; Zürn et al., 2007) oder eine hohe Problemdefinitionsunsicherheit herrscht, wie im Kontext der KI-Einführung (Lemke et al., 2024; Lewallen, 2021). Sie setzt voraus, dass soziale Sachverhalte nicht als naturgegeben, sondern als kontingent, also gestaltbar und offen für alternative Deutungen und Entscheidungen, wahrgenommen werden. Nur wenn politische Entscheidungen nicht als alternativlos erscheinen, sondern als potenziell strittige Optionen unter mehreren, können sie Gegenstand kollektiver Aushandlung werden (Hay, 2007; Zürn, 2013). Diese Wahrnehmung von Kontingenz bringt jedoch nicht nur Entscheidungsoptionen hervor, sondern erzeugt zugleich Rechtfertigungsdruck. Wie Selk (2011) formuliert, wirft das Kontingenzbewusstsein „ein in der politischen Gesellschaft dauerhaft bestehendes Legitimationsproblem auf: Wenn erst einmal eingesehen ist, dass politische Entscheidungen kontingent sind, dann steht jede Entscheidung – auch die Entscheidung, nicht zu entscheiden – unter einem potentiellen Rechtfertigungsdruck“ (S. 195). Somit gilt es etwa für politische Entscheider:innen, zu rechtfertigen, wieso eine KI eingesetzt oder eben auch nicht eingesetzt wird, da nach dem Grundsatz der Kontingenz auch anders hätte entschieden werden können.

Vor diesem Hintergrund wird Politisierung zunehmend als „Prozess [beschrieben], einen sozialen Sachverhalt in zwei Hinsichten zu konstruieren: erstens als *kontingent* und zweitens als *prinzipiell zugänglich für kollektiv bindende Entscheidungen*“ (Ulbricht, 2020, S. 258; Herv. i. O.). Diese ergänzende Definition fasst die bisherigen Überlegungen zusammen, indem sie die doppelte Bewegung betont, die Politisierung ausmacht: die Entnaturalisierung sozialer

Sachverhalte und ihre Überführung in die Arena kollektiver Entscheidungsfindung. Damit wird Politisierung nicht nur als Verschiebung institutioneller Zuständigkeiten oder als Eröffnung öffentlicher Debatten verstanden, sondern als grundlegender Akt sozialer Konstruktion durch die Infragestellung sozialer Normen und Ordnungen, der sowohl die Möglichkeit als auch die Notwendigkeit politischer Entscheidung und ihrer Rechtfertigung hervorbringt. Politisierung wird somit konzipiert als notwendige Bedingung funktionierender Demokratien und einer demokratisch legitimierten Entscheidung über den KI-Einsatz (Ulbricht, 2020). Dabei gilt: „Politisierung geht der Demokratisierung notwendig voraus, freilich gilt einschränkend, dass nicht jede Politisierung zur Demokratisierung führt“ (Zürn, 2013, S. 25).

4.2. Der graduelle demokratische Prozess zwischen Depolitisierung und (Re-)Politisierung

Dies gilt insbesondere, da Politisierung keinen linearen, sondern einen graduellen und reversiblen Prozess bildet, der auch Momente der Depolitisierung umfassen kann, etwa zur Komplexitätsreduktion oder im Sinne strategischer Steuerung (Bates et al., 2014; Bolsen & Druckman, 2015; Hay, 2014; Meiering & Schäfer, 2020; Schimmelfennig, 2021; Selk, 2011; Ulbricht, 2020). Dabei beschreibt Depolitisierung „die Konstruktion eines sozialen Sachverhalts als erstens unveränderbar und/oder zweitens nicht zugänglich oder relevant für kollektiv bindende Entscheidungen“ (Ulbricht, 2020, S. 256). In Anbetracht der Prozesshaftigkeit der Politisierung kann die Einführung von KI in unterschiedlichem Maße politisiert oder depolitisiert sein, je nachdem, ob sie breit öffentlich verhandelt oder etwa technokratisch behandelt wird. Letzteres kann mitunter auf politische Strategien verweisen, durch die Akteur:innen versuchen, kontroverse Entscheidungen, insbesondere im Falle von Hochrisiko-Policies, dem öffentlichen Diskurs zu entziehen (Schimmelfennig, 2021; Zürn, 2013) und somit eine (vermeintlich) konfliktfreie Entscheidungsfindung herbeizuführen (Caramani, 2020). Trotz vielfältiger Risiken, etwa durch unbeabsichtigte oder nicht antizipierbare Folgen, ist die Politisierung technologischer Entwicklungen somit keineswegs selbstverständlich. Wie Saretzki (2001, S. 185) betont, verfolgen verschiedene wirtschaftliche wie auch politische Akteur:innen das Ziel, Technikentwicklung und -einführung möglichst „ungestört“ zu gestalten. Konflikte um technologische Innovationen bleiben daher eher die Ausnahme als die Regel, ist doch das Ziel politischer Entscheidungsträger:innen weniger, Wertkonflikte öffentlich auszutragen, sondern vielmehr, durch problemzentrierte Lösungsrhetorik Zustimmung für bereits definierte politische Pfade zu gewinnen: „The task of policy is to foreground problems; propose solutions to those problems and persuade its audiences (…) of the legitimacy of the proposed problems and solutions“ (Nokkala, 2016, S. 144).

4.2.1. Funktionen und Dysfunktionen der Depolitisierung

Dabei ist Depolitisierung, ebenso wie Politisierung, ein integraler Bestandteil funktionierender Demokratien. Angesichts der Vielzahl potenzieller Konfliktlinien und der umfassenden Komplexität moderner Gesellschaften ist politisches und administratives Handeln nur dann effektiv möglich, wenn nicht alle Themen gleichzeitig politisiert werden. Um soziale Ordnung und Stabilität zu gewährleisten, ist daher eine Komplexitätsreduktion notwendig. Prozesse der (De-)Politisierung übernehmen dabei eine selektierende Funktion: Sie entscheiden darüber, welche Themen in den politischen Raum gelangen und welche (vorübergehend) ausgeblendet bleiben. Auf diese Weise tragen sie wesentlich zur Strukturierung gesellschaftlicher Aufmerksamkeit und zur Herstellung einer handlungsfähigen politischen Ordnung bei. Die Entscheidung darüber, welche Themen ausreichend Salienz und Dringlichkeit besitzen, um öffentlich verhandelt zu werden, ist somit ein zentraler Moment sozialer Ordnungsbildung (Selk, 2011).

Nichtsdestotrotz entstand die Debatte um Depolitisierung zunächst aus einer kritischen Perspektive. Sie ist insbesondere als Reaktion auf neoliberale Transformationsprozesse zu verstehen, welche die Privatisierung öffentlicher Infrastrukturen, eine Schwächung politischer Gestaltungsspielräume sowie eine Zunahme sozialer Ungleichheit begleiteten (Meiering & Schäfer, 2020). In diesem Zusammenhang identifizieren Meiering und Schäfer (2020) verschiedene Stränge der Depolitisierungsforschung, die für die Analyse der Legitimation der KI-Regulierung aufschlussreich sind. Erstens verweist die kritische Governanceforschung auf die Verlagerung von Entscheidungsprozessen in technokratische Verwaltungsorgane, in denen Projekte zunehmend von Wissenschaft, Wirtschaft und Administrator:innen entwickelt und umgesetzt werden (s. z. B. Bertsou & Caramani, 2022; Fawcett et al., 2017). Gerade in Anbetracht des hohen Maßes an Komplexität und Unsicherheit im Kontext der Technologiepolitik versprechen sich politische Entscheider:innen durch eine „Epistemisierung des Politischen“ (Bogner, 2021a) einen Legitimitätsgewinn politischer Entscheidungen. Infolge dieser Verschiebung werden insbesondere kommunale und regionale Politik zunehmend durch wissenschaftlich-technische Expertise ersetzt, was das Spannungsverhältnis zwischen demokratischer Input-Legitimität und effizienter Output-Legitimität auf die Probe stellt, da politische Aushandlungsprozesse durch vermeintlich objektive Lösungen ersetzt werden (Bickerton & Invernizzi Accetti, 2020; Caramani, 2020). Dagegen argumentieren einige Autor:innen, wie zum Beispiel Pettit (2004) oder Rosanvallon (2011), dass die Depolitisierung der politischen Entscheidungsfindung funktional für die demokratische Legitimität sei, indem etwa politische Entscheidungen weniger von wahlpolitischen Einflüssen geprägt und somit unabhängiger und rationaler gefällt werden können. Eng in Verbindung dazu steht der zweite Strang der Depolitisierungsforschung, der

sich auf die Delegation von Entscheidungskompetenzen fokussiert. Indem Entscheidungen zunehmend von Expert:innengremien getroffen werden, „werden politische Konflikte so in zu verwaltende ‚Sachfragen' verwandelt, was die jeweils vertretene Position als alternativlos erscheinen lässt" (Meiering & Schäfer, 2020, S. 13; s. a. Volk, 2013).[6]

Insgesamt manifestiert sich Depolitisierung somit in einer doppelten Verschiebung: dem Rückgang öffentlicher Beteiligung an politischen Entscheidungsprozessen und der Verlagerung politischer Autorität in technokratische Strukturen. Diese Entwicklung kann sich im Kontext von KI als besonders kritisch erweisen, da gerade hier aufgrund der vorhandenen Risiken und weitreichenden Konsequenzen für die Gesellschaft eine komplexe Problemdefinition ausgehandelt werden muss, die nicht allein durch Fachwissen, sondern durch öffentliche Aushandlung zu legitimieren ist (Bartl, 2023; Lemke et al., 2024). Andernfalls sei zu befürchten, dass eine unzureichende Politisierung in eine demokratiegefährdende Dynamik mündet. In der Entpolitisierungsdebatte werden „das Abstreiten eines politischen Entscheidungs(spiel)raumes, die Delegation von Entscheidungskompetenz an technokratische Experten und ein zunehmendes öffentliches Desinteresse an Politik" (Meiering & Schäfer, 2020, S. 14), ausgelöst durch diese Entwicklung, kritisch hervorgehoben.

Bereits die Auslagerung politischer und gesellschaftlicher Entscheidungen auf KI-Anwendungen verdeutlicht dabei die Bedeutung der Depolitisierung durch technokratische Herrschaft im Zuge technologischen Wandels (Ulbricht, 2020). Diese Systeme werden betrachtet „als Übersetzung politischer Konflikte in vermeintlich apolitische, weil scheinbar neutrale und faktenbasierte Entscheidungen und Lösungen" (Katzenbach, 2018, S. 235). Katzenbach kritisiert dieses Verständnis als „funktionalistischen Fehlschluss früherer Technikheorien" (ebd.), da es die soziale Konstruiertheit der Technik ausblendet (siehe Kapitel 6). Somit droht eine Entkoppelung politischer Entscheidungen von gesellschaftlicher Aushandlung und Gestaltbarkeit und damit eine Schwächung der Legitimität der KI-Anwendungen (Bartl, 2023).

4.2.2. Repolitisierung als Konfliktsymptom

Gleichzeitig sind Depolitisierungsprozesse nicht irreversibel. Vielmehr lässt sich beobachten, dass in bestimmten Kontexten eine Repolitisierung stattfindet, also eine Rückkehr politischer Auseinandersetzungen in Themenbereiche, die zuvor als unstrittig oder technisch dargestellt wurden. Dabei zeigen sich gegenwärtige Repolitisierungsprozesse unter anderem in der

[6] Meiering & Schäfer (2020) verweisen darüber hinaus noch auf einen dritten Strang der Depolitisierungsforschung, der im Kontext der KI-Debatte allerdings nur von untergeordneter Rolle ist. So widmet sich dieser Bereich der Transformation des öffentlichen und auch des politischen Raums. Dabei begünstigt die Professionalisierung politischer Parteien eine Priorisierung innerorganisatorischer Logiken (z. B. Personalfragen), während substanzielle gesellschaftliche Konflikte in den Hintergrund treten (Katz & Mair, 1995).

zunehmenden Infragestellung wissenschaftlicher Autorität sowohl durch politische Akteur:innen als auch durch Bürger:innen (Bijker et al., 2009; Gunn, 2019). Damit wird deutlich, dass technologische Expertise nicht mehr unhinterfragt als objektive Entscheidungsgrundlage akzeptiert wird, sondern selbst zum Gegenstand politischer Auseinandersetzung werden kann.

Ein weiteres zentrales Symptom der Repolitisierung liegt in der Entstehung neuer gesellschaftlicher Konfliktlinien im Zuge der Globalisierung, bei denen sich neue Kollektive von Gewinner:innen und Verlierer:innen herausbilden (Grande & Kriesi, 2013). Auch im Bereich digitaler Technologien wie KI treten diese Konfliktlinien zunehmend zutage, etwa thematisch zwischen Effizienzversprechen und Datenschutzbedenken ebenso wie auf Akteurebene zwischen Technologiekonzernen und Staaten oder zwischen Tech-Unternehmen und ihren Mitarbeitenden (Khanal et al., 2025; Ryan et al., 2024; Williamson & Prybutok, 2024). Zudem zeigt sich, dass bestimmte Themen auch wiederholt politisiert werden können. Auf Phasen der Politisierung können Depolitisierung und schließlich erneute Repolitisierung folgen. Dieser zyklische Prozess verdeutlicht, dass die politische Relevanz eines Themas nie dauerhaft gesichert oder ausgeschlossen ist (Hay, 2014; Meiering & Schäfer, 2020). Nicht zuletzt äußert sich Repolitisierung in nationalen und transnationalen Protestbewegungen, die sich explizit gegen die Verfahren und Ergebnisse technokratischer Entscheidungsprozesse wenden (Everson, 2020; Moravcsik, 2002; Rucht, 2013; Thaa, 2013) und am Beispiel der öffentlichen Empörung über das niederländische SyRI-System sichtbar werden. Dabei handelt es sich um eine KI-Anwendung zur Aufdeckung von Sozialleistungsbetrug, welche aufgrund mangelnder Transparenz, fehlender Rechenschaftspflicht und diskriminierender Wirkung sowohl national als auch international Proteste hervorrief (Wieringa, 2023).

4.2.3. Das Politisierungspotenzial Künstlicher Intelligenz

Dabei haben gerade Technologien wie KI (ungeachtet politisch-strategischer Bemühungen um ihre Depolitisierung) das Potenzial, eine intensive Politisierung hervorzurufen, da sie tief in gesellschaftliche, wirtschaftliche und normative Ordnungen eingreifen (Gallego & Kurer, 2022). Eine Politisierung technologischer Entwicklungen erscheint damit einerseits kaum vermeidbar, bedarf jedoch andererseits eines ausgewogenen Maßes, um eine legitime Entscheidungsfindung weiterhin zu ermöglichen (Barry, 2001; Brown, 2015; Pielke, 2007). Diese Notwendigkeit eines ausgewogenen Maßes an Politisierung zeigt sich besonders deutlich in der diskursiven Dimension technologischer Zukunftserwartungen. Dabei sind technologische Entwicklungen – insbesondere solche mit disruptivem Potenzial wie KI – häufig von konkurrierenden narrativen Zukunftsvisionen begleitet, die sowohl mobilisierende als auch entpolitisierende Effekte entfalten können und auf die Kontingenz der politischen Entscheidung verweisen:

„Futures are contingent, they are imagined, fought for, resisted and embraced in the present – in order to draw an imagined future into the real-time now“ (Brown, 2003, S. 17). Der performative Charakter solcher Erwartungen zeigt sich darin, dass sie politische Handlungsrahmen vorstrukturieren, etwa durch das Setzen von Agenden, die Formulierung von Forschungszielen oder die Zuweisung öffentlicher Ressourcen. Indem sich solche Erwartungen verdichten, transformieren sie sich von offenen Versprechen in politische Programme: „When more and more actors share similar expectations, the promises inherent to these expectations are gradually translated into requirements, guidelines and specifications regarding the new technology“ (van Lente et al., 2013, S. 1616). Zugleich birgt der anfängliche Hype, der mit der Einführung neuer Technologien häufig einhergeht, jedoch das Risiko einer übermäßigen Politisierung in Form überhöhter Erwartungen, die bei ausbleibender Erfüllung zu Enttäuschung, Vertrauensverlust und Legitimationskrisen führen können. Überhöhte Zukunftsvisionen können zwar politische und gesellschaftliche Energien mobilisieren, laufen jedoch Gefahr, bei Nichterfüllung delegitimierend zu wirken, sowohl für die Technologie selbst als auch für ihre institutionellen Träger:innen (Brown, 2003). In diesem politisierten Spannungsverhältnis entfaltet sich auch die politische Debatte rund um KI, wie Bareis und Katzenbach (2022) zusammenfassen:

> In sum, AI's political rhetoric about hopes and fears is far from being informative alone. First and foremost, it is *constitutive* as it frames discourses and (im)possibilities; it is *enabling* as it allows political activity (also in the face of a looming threat); it can be *disguising* as it leaves unpleasant societal side-effects and questions about power structures unmentioned and finally also *(de)legitimizing*, bestowing legitimacy upon political leaders or social institutions—or authorizing certain standpoints or disapproving or condemning others (…). (S. 871; Herv. i. O.)

Indem die politische KI-Rhetorik nicht nur Debatten rahmt, sondern auch bestimmt, welche Deutungen legitim erscheinen, nimmt sie Einfluss darauf, auf welcher moralischen Grundlage politische Entscheidungen gerechtfertigt werden. Bevor die Rolle dieser Rechtfertigungen im Legitimationsprozess näher betrachtet wird, folgt zunächst eine Differenzierung des Politisierungsbegriffs entlang zentraler Dimensionen.

4.3. Drei Dimensionen der (De-)Politisierung

Zur Erfassung des graduellen Charakters der Politisierung hat sich in der Politikwissenschaft eine dreidimensionale Differenzierung etabliert. Unterschieden werden die Reichweite der beteiligten Akteur:innen, das Kontingenzbewusstsein sowie die Intensität des Konflikts (de

Wilde, 2011; de Wilde et al., 2016; Grande & Kriesi, 2013; Meiering & Schäfer, 2020).[7] Dabei konzipieren Hutter und Grande (2014) das Kontingenzbewusstsein (von den Autoren als Themensalienz erfasst) als notwendige Bedingung für Politisierung, während die Reichweite der beteiligten Akteur:innen und die Intensität des Konflikts als hinreichende Bedingungen den Grad der Politisierung bestimmen.

So spielen erstens, wie bereits dargelegt, die Salienz des Themas und das damit verknüpfte Kontingenzbewusstsein eine zentrale Rolle. Es bezeichnet das Maß, in dem ein Sachverhalt als gestaltbar, also politisch entscheidbar, wahrgenommen wird. Meiering und Schäfer (2020) definieren dies als den Moment, in dem „die wahrgenommene *Entscheidbarkeit* oder *Gestaltbarkeit* einer Thematik zur Disposition [ge]stellt" (S. 19, Herv. i. O.) wird. Findet ein Thema in der Öffentlichkeit nicht statt und wird es als naturgegeben oder technisch alternativlos dargestellt, bleibt es qua Definition depolitisiert. Erst durch die Wahrnehmung, dass politische Alternativen bestehen, etwa im Umgang mit algorithmischer Entscheidungsunterstützung im Bildungsbereich, wird die Grundlage für einen politisierenden Diskurs geschaffen (Ulbricht, 2020). Zweitens gilt im Hinblick auf die Reichweite der Akteur:innen, die sich in einer Debatte positionieren: Je größer und heterogener das Spektrum politischer und zivilgesellschaftlicher Akteur:innen, welches sich öffentlich zu einem Thema äußert, desto eher wird dieses als politisch relevant wahrgenommen. Die Reichweite kann sich auf den institutionellen Status der Akteur:innen (z. B. Parteien, Verbände, NGOs) ebenso beziehen wie auf deren gesellschaftliche Repräsentativität oder mediale Sichtbarkeit (de Wilde et al., 2016; Hutter & Grande, 2014). Drittens ist die Intensität des Konflikts zu berücksichtigen. Sie bezieht sich auf das Ausmaß an inhaltlicher Opposition und Auseinandersetzung, welches ein Thema hervorruft (de Wilde et al., 2016). Themen, die normative Grundsatzfragen aufwerfen, gesellschaftliche Interessen fundamental berühren oder bestehende Machtverhältnisse infrage stellen, zeigen eine hohe Konfliktintensität und sind somit stärker politisiert. So erfasst de Wilde (2011) den Politisierungsprozess auch als „increase in polarization of opinions, interests or values and the extent to which they are publicly advanced towards the process of policy formulation" (S. 560). In dem Maße, wie sich unterschiedliche Werte und Interessen öffentlich artikulieren und auf politische Entscheidungsprozesse einwirken, entsteht somit ein Bedarf nach Rechtfertigung. Politisierung eröffnet somit den Aushandlungsraum politischer Rechtfertigungen als Grundlage öffentlicher Auseinandersetzung und der Legitimität politischer Entscheidungen.

[7] In der Politisierungsdebatte wurden unterschiedliche Eigenschaften der Politisierung benannt, die sich trotz variierender Terminologie im Kern auf die hier genannten Dimensionen zurückführen lassen. So unterscheidet Schattschneider (1957) Intensität, Sichtbarkeit, Richtung und Reichweite, während de Wilde (2011) die Dimensionen der Meinungspolarisierung, Debattenintensität und öffentliche Resonanz hervorhebt.

5. Zur Rolle der Rechtfertigungen in Legitimationsprozessen

Die Notwendigkeit der politischen Rechtfertigung, wie sie in der SdK als konstitutives Moment legitimer Ordnung herausgearbeitet wird, ist gemeinhin als fundamentales Prinzip legitimer Herrschaft in demokratischen Gesellschaften anerkannt. Spätestens mit der Ausübung politischer Macht und der Inanspruchnahme von Souveränität entsteht ein Legitimationsdruck, dem durch Verfahren öffentlicher Rechtfertigung begegnet wird (Boltanski & Thévenot, 1991/2018; de Fine Licht, 2020; Gadinger, 2016; Grüttner, 2013; Kukkonen et al., 2021; Niemann, 2019; Pitkin, 1967; Simmons, 1999). Rechtfertigungen sind dabei nicht bloß kommunikative Begleiterscheinungen politischer Prozesse, sondern konstitutive Elemente von Legitimitätsüberzeugungen: Sie tragen wesentlich zur Akzeptanz von Entscheidungen bei, auch dann, wenn Bürger:innen mit inhaltlichen Ergebnissen nicht übereinstimmen, sie jedoch als nachvollziehbar und normativ vertretbar erachten (McGraw et al., 1995; Pitkin, 1967). In diesem Sinne ermöglicht die Fokussierung auf Rechtfertigungen in der vorliegenden Arbeit, das theoretisch-normativ aufgeladene Konzept der Legitimation in empirisch beobachtbare und analysierbare kommunikative Muster zu übersetzen. Rechtfertigungen fungieren damit als operationalisierte Form von Legitimationspraxis und bilden eine zentrale Schnittstelle zwischen normativer Ordnung und politischer Praxis.

Dabei beschreibt Susen (2017) Rechtfertigungen im weitesten Sinne als „act of providing reasons for the validity, legitimacy and defensibility of (a) an action, (b) a belief and/or (c) a social arrangement" (S. 350). Im Unterschied zu häufig synonym verwendeten, aber dennoch konzeptuell zu differenzierenden Begriffen der Argumentation oder Begründung zielt die Rechtfertigung dabei nicht nur auf kognitives Verständnis oder logische Stringenz, sondern auf normative Anerkennung.[8] Während Begründungen auf deduktiven oder empirischen Schlussfolgerungen beruhen, intendiert Rechtfertigung die Herstellung von normativer Angemessenheit. Sie verweist auf überindividuelle Werte, Regeln oder Pflichten, um Handlungen oder Entscheidungen als „gut", „angemessen" oder „gerecht" erscheinen zu lassen (Henin & Le Métayer, 2021). Oder anders formuliert: Rechtfertigungen setzen normative Referenzrahmen voraus, ohne die eine Bewertung politischer Ergebnisse nicht möglich wäre:

> [J]ustifications are normative and extrinsic in the sense that they depend on a reference according to which the adequacy or appropriateness of the outcomes can be assessed.

[8] Ebenso betont Schmidt (2002b) im Rahmen des diskursiven Institutionalismus, dass die Begründung politischer Entscheidungen sowohl kognitive als auch normative Funktionen erfüllen muss, um legitimierende Funktionen zu erfüllen. Erstere beziehen sich auf die Darstellung wirksamer, glaubwürdiger und umsetzbarer Problemlösungen, Letztere auf die Anschlussfähigkeit an geltende Normen und Werte.

> Indeed, in order to claim that an outcome is good (or adequate, appropriate) it is necessary to refer to an independent definition of what a good outcome is. (Henin & Le Métayer, 2021, S. 1399)

Besonders deutlich wird diese Differenz im Umgang mit funktionalen Erklärungen algorithmischer Entscheidungen. So reichen technische oder funktionale Begründungen folgenreicher Entscheidungen nicht aus, wenn diese tief in gesellschaftliche Prozesse eingreifen. Ein bloßes Verständnis, wie etwa die Entscheidung technisch zustande gekommen ist, genügt somit nicht. Vielmehr bedarf es einer umfassenden, normativ anschlussfähigen Rechtfertigung, die sich an allgemein anerkannten moralischen Prinzipien orientiert (Diaz-Bone & Schrör, 2022; Züger & Asghari, 2023). Solche Rechtfertigungen explizieren moralische Kriterien, die Individuen ihren Bewertungen zugrunde legen, und verschieben damit die Diskussion vom individuellen Nutzen auf kollektive Belange (Anesa et al., 2022; Boltanski & Thévenot, 1991/2018; Graß, 2015).

Zugleich grenzt sich eine Analyse der politischen Rechtfertigung von kommunikationswissenschaftlichen Framing-Analysen ab, wie sie etwa von Köstler und Ossewaarde (2022) bereits im Kontext der deutschen KI-Debatte vorgelegt wurden. Indem eine Rechtfertigungsanalyse zur Identifikation und Rekonstruktion der tieferliegenden moralischen Ordnungsmuster, auf denen politische Legitimation basiert, beiträgt, geht sie über die Beschreibung diskursiver Rahmungen hinaus. Zwar bieten Frames wichtige Hinweise auf Deutungsmuster politischer Kommunikation, bleiben dabei jedoch häufig auf kontextspezifische Narrative beschränkt, die insbesondere die Salienz eines Themas oder einer Deutung hervorheben (Entman, 1993). Die Analyse von Cités hingegen erlaubt eine systematischere Erfassung normativer Bezüge, die über kommunikative Oberflächen hinausreichen und sich in institutionellen Ordnungen, Praktiken und normativen Ansprüchen politischer Akteur:innen verankern. Damit trägt die Cité-Analyse dazu bei, moralische Legitimationsgrundlagen sichtbar zu machen, die in Framing-Ansätzen häufig nur implizit oder fragmentarisch erfasst werden (Reinecke et al., 2017).

Dabei ist jedoch zu beachten, dass die reine Rechtfertigungspraxis nicht automatisch Akzeptanz erzeugt: Divergierende normative Überzeugungen innerhalb der Öffentlichkeit können dazu führen, dass Entscheidungen trotz Rechtfertigungsversuchen als illegitim wahrgenommen werden. Entscheidend ist daher, dass Rechtfertigungen in der SdK nicht nur als affirmative Praktiken der Legitimationsherstellung, sondern auch als Elemente politischer Auseinandersetzung und kritischer Zurückweisung verstanden werden: „Justifications can involve positive ‚arguments', claims, or position statements, but might also be critical ‚denunciations' of opposing views in the dynamics of public disputes" (Thévenot et al., 2000, S. 237). Insofern ist

Rechtfertigung konflikthaft verfasst. Sie verweist auf konkurrierende normative Geltungsansprüche, deren Verhältnis zueinander nicht apriorisch entschieden ist, sondern in öffentlichen, politisierten Auseinandersetzungen ausgehandelt werden muss, um eine demokratisch legitimierte Entscheidung zu ermöglichen: „Justification can be considered a linchpin of normative controversy, because it combines the role of norms, legitimation, and politicisation“ (Niemann, 2019, S. 65). Während bereits argumentiert wurde, welche Rolle Politisierung und Rechtfertigungen im Legitimationsprozess der politischen Entscheidung über die KI-Einführung spielen, liegt der weitere Fokus nun zunächst auf der moralischen Grundlage dieser Rechtfertigungen, bevor anschließend die Notwendigkeit der Kritik im Sinne eines konstitutiven Wechselspiels in Legitimationsprozessen dargelegt wird.

5.1. Cités als moralische Grundlage der Rechtfertigung und Kritik

Die moralische Fundierung einer Rechtfertigung fassen Boltanski und Thévenot (1991/2018) im Konzept der *Cité*, einem Begriff, der in der deutschsprachigen Rezeption auch als Rechtfertigungsordnung, Konvention, Polis oder Gemeinwesen übersetzt wurde.[9] Cités bezeichnen unterschiedliche normative Ordnungen innerhalb einer Gesellschaft, die in sozialen Auseinandersetzungen von zentraler Bedeutung sind, da sie als moralisch legitimierte Bezugspunkte für das eigene Handeln, Entscheiden und Fordern herangezogen werden (Bogusz, 2010; Diaz-Bone, 2022; Forst, 2015). Sie bieten normative Referenzrahmen, welche es Akteur:innen ermöglichen, partikulare Interessen in allgemein akzeptable Rechtfertigungsnarrative zu überführen. Dabei sind Cités nicht selbst als Rechtfertigungen zu verstehen, sondern als Äquivalenzprinzipien, die sowohl der Formulierung von Rechtfertigungen als auch von Kritik eines politischen Standpunkts zugrunde liegen können. Um terminologische Missverständnisse zu vermeiden, wird im Folgenden auf eine Übersetzung verzichtet und der französische Originalbegriff „Cité“ beibehalten.

In einer politisierten Welt, die durch Kontingenz und Ungewissheit geprägt ist, versteht die SdK den Rückgriff auf einzelne Cités als eine Form der Komplexitätsreduktion (Bogusz, 2010; Boltanski & Thévenot, 1991/2018; Niemann, 2019). Akteur:innen greifen dabei auf diese moralischen Ordnungsprinzipien als Koordinationslogiken zurück, um Handlungen und Positionen zu bewerten, zu legitimieren oder zu kritisieren und so eine koordinierte Bezugnahme im sozialen Raum zu ermöglichen (Diaz-Bone, 2022; Honneth, 2008; Nachtwey & Seidl, 2021). So kann beispielsweise der Einsatz eines KI-basierten Lernempfehlungssystems mit dem

[9] Eine jede Cité wird durch sechs konstituierende Axiome gekennzeichnet, welche unter anderem auf der Anerkennung des ‚gemeinsamen Menschseins‘ und prinzipiell gleichen Zugangschancen basieren. Gleichzeitig berücksichtigen diese Axiome jedoch auch die Unterschiedlichkeit von Individuen, die es unter Heranziehung des zugrunde liegenden Gemeinwohls zu rechtfertigen gilt (Boltanski & Thévenot, 1991/2018, S. 108 ff.).

Verweis auf Effizienzgewinne im Lernprozess gerechtfertigt werden. Dabei ermöglicht der Bezug auf Effizienz es anderen Akteur:innen, diesen aufzugreifen oder zu kritisieren, da dieser auf einer gesellschaftlichen Konvention beruht, die allgemein verständlich ist, auch wenn ihr nicht zwangsläufig zugestimmt wird. In Anlehnung an ein normatives Verständnis des „Allgemeinen" (Boltanski & Thévenot, 1991/2018) eröffnen Cités so die Möglichkeit, über geteilte Werte und Gerechtigkeitsvorstellungen tragfähige Kompromisse in demokratischen Aushandlungsprozessen zu erzielen:

> While totalitarian regimes can try to suppress an imperative of justification, democratic societies can only preserve their values by enabling public discourses of justification. That is, ordinary actors continually implement critique and justification and rise up against domination by constructing new interpretations of reality. When actors invoke justice demands, they see each other in a world of common humanity and are principally able to agree to a compromise. (Gadinger, 2016, S. 192 f.)

Politische Akteur:innen berufen sich auf bestimmte Cités, weil sie davon überzeugt sind, dass diese im jeweiligen Kontext eine angemessene Grundlage für legitimes Handeln bieten (Anesa et al., 2022; Graß, 2015). Der moralische Gehalt einer Cité ergibt sich dabei aus der ihr zugrunde liegenden Gerechtigkeitsvorstellung, die sich stets auf das Gemeinwohl bezieht (Hartmann, 2008). Dies umfasst nicht nur ethisch-normative Moralvorstellungen im alltagsweltlichen Sinne, sondern auch situativ begründete Orientierungen wie etwa Effizienz, Planungssicherheit oder Wettbewerb. Der in der Arbeit verwendete Begriff der moralischen Grundlage ist daher als Bezug auf gesellschaftliche Konventionen zu verstehen, die den Ordnungsrahmen gesellschaftlichen Miteinanders bestimmen: „Dieser (die eigene Handlungsweise rechtfertigende) Bezug auf das ‚Gemeinwohl' ist virtuell in den Situationen eben vermittels der Konventionen so für Akteure präsent, dass letztere die Koordination bewusst oder unbewusst auf ein Gemeinwohl als (anzustrebendes) Ziel beziehen können" (Diaz-Bone, 2022, S. 476).

Ursprünglich differenzieren Boltanski und Thévenot (1991/2018) sechs Cités, die als moralisch fundierte Koordinationslogiken in gesellschaftlichen Auseinandersetzungen herangezogen werden können: die Cité des Marktes, der Industrie, der Inspiration, der (öffentlichen) Meinung, die häusliche und die staatsbürgerliche Cité. Diese Ordnungsmuster sind jedoch nicht als statisches Analyseraster zu verstehen, sondern als „Rekonstruktion der *empirisch* vorhandenen moralischen Ordnungen" (Diaz-Bone, 2018, S. 145, Herv. i. O.)[10], die von prinzipieller Offenheit

[10] Boltanski und Thévenot rekonstruieren die von ihnen orginär benannten Cités selbst empirisch auf Grundlage einflussreicher Managementliteratur (s. a. Diaz-Bone, 2022).

und Veränderbarkeit geprägt sind (Graß, 2015). Ihre Geltung ist nicht dauerhaft garantiert, sondern muss im sozialen Handeln stets neu hervorgebracht und bestätigt werden (Bogusz, 2010). So sind im Zuge gesellschaftlicher Transformationen neue Cités identifiziert worden, etwa eine projektbasierte Cité (Boltanski & Chiapello, 2003), eine ökologische Cité (Lafaye & Thévenot, 1993; Thévenot et al., 2000) sowie, jüngeren technopolitischen Diskursen folgend, eine Cité der Solution, in der gesellschaftliche Probleme primär als technische definiert und technokratisch bearbeitet werden (Nachtwey & Seidl, 2021, 2023).[11]

Dabei sind Cités weder mit sozialen Milieus zu verwechseln noch auf bestimmte gesellschaftliche Sphären beschränkt. Akteur:innen argumentieren nicht ausschließlich aus der Perspektive einer Cité, sondern können in unterschiedlichen Kontexten auf verschiedene moralische Logiken Bezug nehmen (Bogusz, 2010; Boltanski & Thévenot, 1991/2018; Hartmann, 2008). Sie müssen daher „als *entdifferenziert* gedacht werden (…). Es gibt je eine Gleichzeitigkeit von Koordinations- und Bewertungslogiken, anstatt dass jede Konvention eine spezifische soziale Sphäre strukturiert“ (Diaz-Bone, 2022, S. 489; Herv. i. O.). Gerade diese Gleichzeitigkeit eröffnet die Möglichkeit von Konflikten, wenn in einer Situation konkurrierende Rechtfertigungslogiken aufeinandertreffen (Boltanski & Thévenot, 1991/2018; siehe Abschnitt 5.2.2).

Im Folgenden werden die zentralen Cités, wie sie im theoretischen Rahmen der SdK konzipiert sind, systematisch dargestellt. Dabei gewinnen im Zuge des New Public Managements im Bildungssektor insbesondere die Cités des Staatsbürgers, des Marktes und der Industrie an Bedeutung, weshalb diese zuerst dargelegt werden (Turnbull et al., 2024). Ziel ist es, ihre jeweilige moralische Grundlage und ihre Bewertungslogik im Kontext technologiepolitischer Auseinandersetzungen zu erläutern. Dies bildet die Basis, um an späterer Stelle die politische Rechtfertigung und Kritik empirisch zu analysieren und resultierende Kompromisse wie auch Konfliktlinien der bildungspolitischen Debatte über KI zu identifizieren sowie in Hinblick auf ihren Beitrag zur Legitimation der Entscheidung der KI-Einführung zu diskutieren.

5.1.1. Cité des Staatsbürgers

In der staatsbürgerlichen, bürgerweltlichen oder auch zivilgesellschaftlichen Cité stehen das Kollektiv und das alltagsweltliche Verständnis des Gemeinwohls der Gesellschaft[12] im Vordergrund. Demnach gilt es, individuelle Interessen zugunsten von Solidarität und kollektiven

[11] Auch die im Zuge der durch die Corona-Pandemie angestoßenen gesundheitspolitischen Debatten hervorgetretene *Cité* der Vitalität (Sharon, 2018; Siffels, 2021) sei der Vollständigkeit halber erwähnt, wenngleich sie im Folgenden nicht näher behandelt wird.

[12] Gemeinwohl wird an dieser Stelle verstanden als das der Allgemeinheit, also der Gesellschaft, dienende Wohl und grenzt sich insofern von dem von Boltanski & Thévenot (1991/2018) verwendeten Gemeinwohlbegriff ab, der den angestrebten Idealzustand einer jeden einzelnen Cité beschreibt und somit multiple Formen annimmt bzw. unterschiedlichen gesellschaftlichen Teilbereichen dienen kann.

Interessen zurückzustellen (Forst, 2015; Gadinger & Yildiz, 2012). In der Konsequenz soll Individualismus vermieden werden. Somit spielen Teilhabe und Mitbestimmung wichtige Rollen in der demokratischen Entscheidungsfindung, während Willkür verurteilt wird (Bogusz, 2010). Bildungspolitisches Ziel ist es in dieser Perspektive, Schüler:innen einem humanistischen Bildungsideal folgend zu mündigen Bürger:innen auszubilden (Frei et al., 2023; Imdorf, 2011; Imdorf & Leemann, 2023). Auch die häufig mit algorithmischen Entscheidungen in Verbindung gebrachte Kritik der (Verschärfung von) Ungleichheit und Diskriminierung tritt in dieser Cité besonders deutlich zutage, wobei Fragen der Chancen- und Bildungsgerechtigkeit aufgeworfen werden (Graß, 2015; Imdorf, 2011; Imdorf & Leemann, 2023; Leemann, 2014; Peter, 2015). Eine besondere Rolle kommt außerdem der Legalität, verstanden als Rechtmäßigkeit, zu. Der Gemeinwille soll möglichst in Form von Gesetzen und Beschlüssen festgeschrieben werden, sodass individuelle Abweichungen und die Verfolgung egoistischer Interessen erschwert werden (Knoll, 2012). Außerdem wird eine demokratische Repräsentativität des Kollektivs gefordert, die ebenfalls festen Gesetzesgrundlagen unterliegt. So sind Repräsentant:innen gegenüber dem Kollektiv stets rechenschaftspflichtig (Boltanski & Thévenot, 1991/2018).

In diesem Sinne erscheinen aus Sicht der Cité des Staatsbürgers sowohl die Forderung nach gesetzlich verankerter Regulierung als auch nach umfassender Rechenschaftspflicht algorithmischer Systeme erwartbar. Ebenso anschlussfähig sind die bereits diskutierten Prinzipien einer menschenzentrierten KI-Entwicklung. Der Einsatz von KI wäre in dieser Logik nur dann legitim, wenn er diskriminierungsfrei und am Gemeinwohl orientiert erfolgt oder etwa demokratische Teilhabe stärkt.

5.1.2. Cité des Marktes

Im Gegensatz zur staatsbürgerlichen Cité kommt dem Individualismus in der Cité des Marktes durchaus eine Bedeutung zu. So wird etwa bildungspolitisch eine Gleichbehandlung als hinderlich betrachtet, da es gelte, individuelle Stärken zu fördern und Lernende so optimal für den Arbeitsmarkt auszubilden (Frei et al., 2023; Imdorf & Leemann, 2023). Dabei ist diese Cité besonders durch Konkurrenz geprägt, die durch die ökonomische Grundannahme der Begierde nach knappen Ressourcen entsteht und somit durch Angebot und Nachfrage geprägt wird. Demnach treten zudem Wettbewerbe und Rivalität zutage, die Ausdruck des Strebens nach Reichtum und Erfolg sind (Bogusz, 2010; Boltanski & Thévenot, 1991/2018; Imdorf, 2011). In der Praxis äußern sich marktwirtschaftliche Rechtfertigungen daher häufig in einer starken Wettbewerbs- und Kundenorientierung (Nachtwey & Seidl, 2021). So können etwa Bildungseinrichtungen bestrebt sein, sich durch eine individuelle Profilbildung von anderen Einrichtungen abzuheben (Graß, 2015; Imdorf, 2011; Imdorf & Leemann, 2023), wobei ein Mittel zur

Profilbildung auch der KI-Einsatz sein könnte. Zugleich bedingt das Wettbewerbsstreben, dass Akteur:innen entweder in die Kategorie der Gewinner:innen oder in die der Verlierer:innen fallen (Knoll, 2012).

Für die Rechtfertigung des KI-Einsatzes kann somit angenommen werden, dass insbesondere die internationale Wettbewerbsfähigkeit unterstrichen wird. Somit sind Vergleiche mit anderen Ländern zu erwarten, aber auch Bestrebungen, den Profit zu erhöhen und sich von der internationalen Konkurrenz abzuheben.

5.1.3. Cité der Industrie

Obwohl gewisse Parallelen zwischen der Cité der Industrie und der des Marktes zu liegen scheinen, betonen Boltanski und Thévenot (1991/2018) die Andersartigkeit und kritisieren im Zuge dessen gleichsam gegenteilige Auffassungen in der Soziologie. Denn nicht das Streben nach Reichtum ist hier die Antriebsfeder des Handelns, sondern die Innovation, Funktionalität und Effizienz. Ein bildungspolitisches Ziel kann somit sein, die Effizienz schulischer Bildung durch den KI-Einsatz zu erhöhen (Horvath et al., 2025). Um kontinuierlichen Fortschritt und Optimierung, Leistung und Produktivität zu erreichen, bedarf es jedoch auch eines gewissen Maßes an Stabilität und Kontrolle (Bogusz, 2010). In diesem Sinne ist die Cité der Industrie auch durch eine Metrisierung in Form von Kennziffern und Kriterien geprägt, die die Formulierung klarer Zielvorgaben gestatten (Frei et al., 2023; Graß, 2015; Imdorf, 2011; Imdorf & Leemann, 2023). Daraus entstünde eine gewisse zeitliche Äquivalenz, die sich besonders in Tools widerspiegelt, „die sich wie etwa *Programme*, *Pläne* oder *Budgets* dazu eignen, die Zukunft zu managen“, so Boltanski und Thévenot (1991/2018, S. 282, Herv. i. O.). Wenngleich der technologische und wissenschaftliche Fortschritt nicht uneingeschränkt der industriellen Cité vorbehalten ist, wird diese doch traditionell hier verortet (Knoll, 2012). So wird technologischer Fortschritt im Sinne eines Innovationsversprechens aus Sicht dieser Cité als unabdingbar gesehen. Demnach können etwa die häufig mit dem KI-Einsatz verknüpften Versprechungen der Effizienzgewinne und der Standardisierung (siehe Abschnitt 2.2) typischerweise hier verortet werden.

5.1.4. Cité der Inspiration

Die Cité der Inspiration, verschiedentlich auch als Welt der Erleuchtung bezeichnet (Hartmann, 2008), bildet einen abstrakten, übernatürlichen Bezugspunkt, der sich am gottgegebenen orientiert und insofern religiös motiviert scheint (Forst, 2015). Demnach werden Hochmut und Habgier ebenso wie das Streben nach Ruhm gänzlich abgelehnt und anstelle dessen wird Demut als Tugend geschätzt. Zur Erreichung der Inspiration, die von Boltanski und Thévenot (1991/2018) gewissermaßen als esoterischer und berauschender Zustand beschrieben wird, bedarf es den Autoren zufolge asketischer Enthaltsamkeit, insofern als dass öffentliche Aufmerksamkeit der

Gedankenfreiheit im Wege stehe und die Anerkennung und Meinung anderer irrelevant oder gar hinderlich erscheine. In der moderneren Fassung kann die Cité der Inspiration vornehmlich als Welt der Künstler:innen begriffen werden, welche Raum für Kreativität und Selbstverwirklichung bietet und großen Wert auf Ästhetik legt (Frei et al., 2023; Gadinger & Yildiz, 2012; Imdorf & Leemann, 2023). Das Streben nach Originalität mündet in einem Enthusiasmus, der eine rationale Haltung vermissen lässt und gar erfordert, diese wissentlich aufzugeben (Boltanski & Thévenot, 1991/2018). Aufgrund dieser kreativen Ausrichtung erscheint die Cité der Inspiration gewissermaßen phantastisch und sehr impulsiv und ist somit die fluktuativste und unbeständigste Cité (Bogusz, 2010).

Übertragen auf die konkrete Betrachtung der Entwicklung und Rechtfertigung moderner Technologien spiegelt sich die Cité der Inspiration in der Bewunderung von Genialität und der „zuweilen guruartige[n] Verehrung ‚gnadenbegabter' Entrepreneure (man denke nur an Steve Jobs)" (Nachtwey & Seidl, 2021, S. 248) wider. Vorstellbar ist jedoch auch, dass aus dieser Perspektive heraus KI selbst gewissermaßen sakralisiert wird, indem sie als Heilsbringer zur Erreichung höherer Zwecke, wie beispielsweise der Freiheit, symbolisiert wird.

5.1.5. Cité des Hauses

Durch Wertschätzung von Gewohnheit und Traditionen, Vertrauen, Verantwortungsgefühl und Respekt sowie die Ablehnung von Neid, Schwätzertum und Verrat zeichnet sich die häusliche Cité aus (Bogusz, 2010; Boltanski & Thévenot, 1991/2018). Gleichzeitig werden die Akteurskonstellationen in dieser Welt stark durch eine als natürlich gegeben angenommene Hierarchie geprägt, wie sie in familiären Kontexten zu finden ist und persönliche Beziehungen und mitunter autoritäre Abhängigkeitsverhältnisse erzeugt (Forst, 2015; Frei et al., 2023; Gadinger & Yildiz, 2012; Hartmann, 2008; Hildebrandt, 2018; Imdorf, 2011; Leemann, 2014). Dabei muss die Rechtfertigung aus Sicht der häuslichen Cité jedoch nicht ausschließlich familiär geprägt sein, sondern kann sich auch beispielsweise in den Hierarchiestrukturen tradierter Unternehmen oder Bildungsinstitutionen wiederfinden. Im Gegensatz zur Cité der Inspiration gilt daher nicht das „Geniale" oder das „Phantastische" als erstrebenswert, sondern das Tradierte. Häufig wird dabei Bezug auf das Handeln solcher Akteur:innen genommen, die sich bereits verdient gemacht haben und insofern in der Hierarchie oben stehen (Boltanski & Thévenot, 1991/2018).

Vor dem Hintergrund des Traditionsbewusstseins ist davon auszugehen, dass eine KI-Einführung kritisch gesehen wird und insbesondere dann abgelehnt wird, wenn sie droht, bestehende Hierarchien und Rollenbilder zu gefährden. So könnte etwa eine Bestätigung des Status quo als Rechtfertigung aus Sicht der häuslichen Cité herangezogen werden.

5.1.6. Cité der Meinung

Ähnlich wie die Cité der Inspiration ist auch die Cité der Meinung stark fluktuativ. Wiederum in Abgrenzung zur inspirierten Welt gelten hier jedoch Ruhm und Anerkennung als erstrebenswerte Ziele. Die Reputation und das Renommee des Einzelnen oder einer Institution stehen im Fokus der Rechtfertigung (Gadinger & Yildiz, 2012; Hartmann, 2008; Imdorf & Leemann, 2023). Die Cité wird maßgeblich durch die in Medien vermittelte öffentliche Meinung mitgeprägt, sodass Rechtfertigungen auf Basis (der Annahme über) Mehrheitsmeinungen konstruiert werden und den Aussagen sogenannter Meinungsführer:innen besondere Beachtung geschenkt wird. Eigene Interessen werden insofern hintenangestellt, als dass das Streben nach öffentlicher Anerkennung und Wertschätzung in einem Abhängigkeitsverhältnis zu der Meinung anderer mündet (Boltanski & Chiapello, 2001; Boltanski & Thévenot, 1991/2018).

Plausible Rechtfertigungen für oder gegen den Einsatz von KI aus Sicht dieser Cité könnten daher beispielsweise von Politiker:innen geäußert werden, die sich an einem größtmöglichen Wähler:innenpotenzial orientieren. Auch der Bezug zu anderen berühmten Persönlichkeiten oder „erfolgreichen" Staaten kann als Indikator aufgefasst werden (Boltanski & Thévenot, 1991/2018).

5.1.7. Cité des Projekts

In „Der neue Geist des Kapitalismus" erläutern Boltanski und Chiapello (2003) die Entstehung einer projektbasierten Cité, welche maßgeblich durch neoliberale Entwicklungen und Kritik am Kapitalismus geprägt sei. Somit rückt die empirische Beobachtung gesellschaftlichen Wandels zunehmend in den Fokus der SdK. Maßgeblich durch die von der 68er-Bewegung vorgebrachte Kritik[13] an Ungleichheit, autoritären Strukturen, starren Hierarchien, Unterdrückung und der „Uniformierung in einer Massengesellschaft" (Boltanski & Chiapello, 2001, S. 468) geprägt, sei ein *kapitalistischer Geist* entstanden, der eine neue Rechtfertigung als Legitimitätsgrundlage gegenwärtiger Strukturen schaffe (Blättler, 2008; Diaz-Bone, 2022). Der an Gedanken Webers (1905/1993) und Sombarts (1902) anknüpfende kapitalistische Geist fungiert dabei „als motivationale und legitimatorische Bedingung der Möglichkeit für die lebenspraktische Durchsetzung einer rationalen Wirtschaftsgesinnung" (Nachtwey & Seidl, 2021, S. 241). Dies beschreibt ein ideologisches Konstrukt, welches die Verfolgung eines höheren Sinns erfasst und über die Kapitalakkumulation hinausgeht, um das Engagement insbesondere der Arbeitnehmer:innen zu

[13] Boltanski & Chiapello (2001) unterscheiden zwischen zwei Formen der Kapitalismuskritik: der Sozialkritik auf der einen Seite und der Künstlerkritik auf der anderen Seite. Während die Sozialkritik die durch den Kapitalismus befeuerten Ungleichheiten und Ausbeutung sowie die Förderung von Individualismus anstelle von Solidarität bemängelt, kritisiert die Künstlerkritik die Unterdrückung durch die Macht des Marktes und die Gleichmachung und fordert Autonomie, Authentizität und Freiheit (S. 468).

erklären, die in einem kapitalistischen System nicht zu den Profiteur:innen gehören (Boltanski & Chiapello, 2003; Doezema & Frahm, 2023).

Zentral sind dabei der Abbau starrer Hierarchien auf der einen Seite und die Entwicklung netzwerkartiger Arbeitsstrukturen auf der anderen Seite (Bogusz, 2010). Im Gegensatz zur häuslichen Cité wird das Tradierte abgelehnt und Veränderung stellt einen Wert an sich dar. Insofern wird auch kein übergeordnetes Ziel verfolgt, sondern vielmehr eine Reihe von Projekten, die nach Abschluss durch neue ersetzt werden. Mitgestaltung und Teamarbeit spielen eine wichtige Rolle, aber auch Autonomie, Freiheit und Authentizität (Boltanski & Chiapello, 2001). Rechtfertigungen aus dem Blickwinkel der projektbasierten Cité rücken also einzelne Probleme und die Erarbeitung ihrer Lösung in den Fokus, ohne dabei nach allgemeiner Gültigkeit zu streben. Gadinger und Yildiz (2012) analysieren auf Basis der politischen Debatte über die Finanzkrise in Deutschland Rechtfertigungspraktiken, die die Notwendigkeit international-kooperativer Lösungswege in den Vordergrund rücken. Entsprechende Vorstöße sind auch im Kontext der KI-Entwicklung vorstellbar. Rechtfertigungen des Einsatzes von KI könnten in diesem Sinne anwendungsbezogen kontextualisiert werden und nicht etwa darauf bedacht sein, eine allgemeingültige „vertrauenswürdige KI“ (Europäische Kommission, 2019) zu garantieren. Stattdessen sollten sie in einer spezifischen Situation dazu dienen, einen spezifischen Zweck zu erreichen. Zur Erreichung dieses Zwecks kann es vonnöten sein, Expertise ‚von außen‘ einzuholen und in diesem Sinne ein neues Netzwerk zu bilden. Ist das Ziel jedoch erfüllt, können neue Projekte, das heißt neue KI-Lösungen für andere Anwendungsfelder, betrachtet werden, die wiederum den Aufbau neuer Netzwerke erforderlich machen können.

5.1.8. Cité der Ökologie

Auch die Cité der Ökologie oder auch grüne Cité ist mit der Reaktion auf globale Umweltbewegungen gewissermaßen als Folge sozialen Wandels und gesellschaftlicher Kritik zu betrachten (Lafaye & Thévenot, 1993; Thévenot et al., 2000) und gewinnt somit in heutigen Debatten an Aktualität. Analog zu Forderungen gegenwärtiger Klimabewegungen wie *Fridays for Future* stehen Werte der Nachhaltigkeit und Erneuerbarkeit im Vordergrund. Aber auch das Wohlbefinden und die Gesundheit werden thematisiert. Der Blick ist dabei nach vorne gerichtet und potenzielle Szenarien und Auswirkungen auf zukünftige Generationen werden berücksichtigt (Thévenot et al., 2000). Rechtfertigungen, die die Einzigartigkeit der Natur ignorieren und Konsum oder Verschwendung fördern, werden nicht toleriert (Hildebrandt, 2018). Gleichzeitig wird die Cité der Ökologie jedoch auch mit Technikforschung und Wissenschaft verbunden, indem beispielsweise immer wieder Grenzwerte gesetzt werden oder Treibhausgasemissionen verringert werden sollen (Knoll, 2012, S. 70).

Im Kontext von KI lassen sich daraus Rechtfertigungen ableiten, die auf den potenziellen Beitrag technologischer Systeme zu ökologischer Nachhaltigkeit verweisen, etwa durch ressourcenschonende Optimierung von Produktionsprozessen, intelligente Steuerung von Energieverbrauch oder die Unterstützung ökologischer Forschung. Der Einsatz von KI wäre aus Sicht dieser Cité nur dann legitim, wenn er nachweislich zur Bewahrung natürlicher Lebensgrundlagen beiträgt, ökologische Belastungen reduziert und langfristig mit planetaren Grenzen vereinbar bleibt.

5.1.9. Cité der Solution

Die Cité der Solution greift die von Morozov (2013) kritisierte Auffassung des Solutionismus auf, dass ein jedes Problem als technisches Problem und im Besonderen als technisch zu lösendes Problem zu verstehen sei (für eine vertiefte Auseinandersetzung siehe Abschnitte 6.1 sowie 6.4.4). Wenngleich Morozov dieses Phänomen nicht als gänzlich neues betrachtet, betont er dennoch den Bedeutungsgewinn des Solutionismus im Zusammenhang mit dem digitalen Wandel und dem Machtgewinn großer Internetkonzerne sowie insbesondere der Entstehung des Silicon Valleys. So sei der gegenwärtige Kapitalismus beziehungsweise die Kapitalakkumulation untrennbar mit der Entwicklung und dem Einsatz neuer Technologien verbunden (Morozov, 2016). Nachtwey und Seidl (2021) fassen auf dieser Basis zusammen, dass die Entstehung eines digitalkapitalistischen Geistes zu beobachten sei, der sich gewissermaßen als Weiterentwicklung des kapitalistischen Geistes der projektbasierten Cité konstituiere. Der Geist des digitalen Kapitalismus reagiere

> (…) einerseits auf Verschiebungen des Akkumulationsregimes in Form der Herausbildung des digitalen Kapitalismus; andererseits auf Legitimationsprobleme eines moralisch und politisch entfesselten Finanzmarkt-Kapitalismus, wie sie – seit der weltweiten Finanz- und Wirtschaftskrise mit zunehmender Schärfe – vor allem von Seiten der Sozialkritik artikuliert werden. (Nachtwey & Seidl, 2021, S. 242)

Auf Basis dieser Kritik knüpft das solutionistische Verständnis auch an das technokratische Bewusstsein an, das auf der einen Seite technisch-wissenschaftliche Expertise in den politischen Entscheidungsraum einbindet, auf der anderen Seite jedoch Entpolitisierungstendenzen aufweist, indem gesellschaftliche Probleme und politisch-demokratische Entscheidungsbefugnis durch technologische Antworten substituiert werden (Nachtwey & Seidl, 2021).

Vor dem Hintergrund dieser geradezu technikverherrlichenden Weltanschauung kann die Cité der Solution in besonderem Maße als Rechtfertigungsnarrativ der Einführung von KI-Systemen herangezogen werden. Anknüpfend an Morozov (2013) bedarf es demnach zuerst einer

Problemartikulation, unabhängig von der tatsächlichen gesellschaftlichen Bedeutung des Problems, und anschließend der Vorstellung der technologischen Lösbarkeit. Dabei wird diese Lösung entweder als einzig mögliche Option oder zumindest als die bestmögliche Option dargestellt und bietet daher aus Sicht der Solutionist:innen die legitimatorische Grundlage der KI-Implementierung. Von besonderer Bedeutung ist dabei die Auffassung, zur Weltverbesserung beizutragen. Dabei ist Stillstand negativ konnotiert und Veränderung stets positiv. Institutionen und Regularien gelten gleichzeitig als Hindernis für die Optimierungsbestrebungen, weil sie der Innovation im Wege stehen (Nachtwey & Seidl, 2021).

5.2. Kompromisse als Ziel einer legitimen Entscheidung

Während diese Cités die potenzielle moralische Grundlage für die Rechtfertigung des KI-Einsatzes bilden, können sie im Sinne eines konflikthaften Begriffsverständnisses zugleich auch als Grundlage für Kritik fungieren (Thévenot et al., 2000). Sie können somit auch zur Rechtfertigung der Ablehnung des KI-Einsatzes herangezogen werden. Dabei bedürfen legitime Entscheidungen je nach Lesart eines Moments der Unruhe (Boltanski, 2008), der Unsicherheit (Gadinger, 2016), des Konflikts (Forst, 2018) oder der Kontroverse (Niemann, 2019). Denn all diesen Momenten gemein ist, dass bestehende Entscheidungen und soziale Ordnungen hinterfragt werden und im Zuge dessen ausgehandelt wird, welcher Geltungsanspruch einer bestimmten Cité in der jeweiligen Situation als legitim anerkannt werden kann: „Controversy becomes a moment for testing competing justifications in order to determine which principle of worth represents a shared understanding of how to evaluate the situation normatively“ (Niemann, 2019, S. 6). Dabei ist es das zentrale Anliegen der pragmatischen Soziologie Boltanskis (2011), Momente der Unsicherheit hervorzuheben, die die soziale Realität herausfordern. Diese Momente lösen Streitigkeiten innerhalb einzelner Cités oder Konflikte zwischen verschiedenen Cités über die (Re-)Konstruktion der Wirklichkeit aus (Bogusz, 2010; Boltanski & Thévenot, 1991/2018). Die daraus entstehenden Kontroversen werden im Kontext der Entwicklung und Einführung neuer Technologien insbesondere durch die Unsicherheit darüber ausgelöst, was KI-Anwendungen tatsächlich zu leisten imstande sind und welche Folgen sie haben. So kann die KI-Einführung als kritischer Moment im Bildungswesen verstanden werden, der die bestehende Ordnung infrage stellt und Interpretations- und Handlungsspielräume bezüglich geeigneter politischer Maßnahmen eröffnet (Obendiek & Seidl, 2023). Zugleich erfordert dies die Aushandlung von Kompromissen, um etablierte Routinen und Normen neu zu verhandeln (Leemann, 2014).

Während diese Konfliktsituationen den Untersuchungsgegenstand der SdK bilden, liegt das weitere Erkenntnisinteresse in der Untersuchung der (Wieder-)Herstellung der Einigkeit und

folglich der Stabilisierung sozialer Ordnung. Für Boltanski und Thévenot (1991/2018) münden diese durch Ungewissheit ausgelösten Streitigkeiten in sogenannten „Prüfungen", die als Prozess zu verstehen sind, der „die Uneinigkeit beendet, indem sie wieder eine richtige Anordnung von Personen und Objekten vornimmt" (S. 191), die soziale Ordnung also wiederherstellt. Sie dienen demnach der Aushandlung einer gemeinsamen Realitätskonstruktion (Boltanski & Thévenot, 1991/2018; Smith Ochoa et al., 2021), sodass eine Verständigung stattfindet, welche Cité ein gemeinsames Verständnis der Situation ermöglicht und so eine legitime Entscheidung herbeiführen kann. Im Zuge dieser Prüfungs- und Einigungsprozesse schließen die beteiligten Akteur:innen Kompromisse, um trotz konkurrierender Cité-Ansprüche eine für alle Beteiligten akzeptable Lösung zu erzielen. Kompromisse ermöglichen es, konkurrierende moralische Prinzipien in einer pragmatischen Balance zu vereinen und auf diese Weise eine geteilte Legitimität politischer Entscheidungen herzustellen, ohne bestehende Differenzen vollständig aufzulösen (Bogusz, 2013; Niemann, 2019). Somit gelten Prüfungen auch als „Einigungsprinzip par excellence" (Bogusz, 2010, S. 51), die dazu dienen, eine legitime Entscheidung zu ermöglichen.

Demgemäß sind Kompromisse als Ausdruck nivellierender Aushandlungsprozesse zu verstehen, in denen unterschiedliche moralische Orientierungen nicht gegeneinander ausgespielt, sondern wechselseitig ergänzt werden. Legitime Entscheidungen entstehen dabei entweder durch die Anerkennung der Gültigkeit einer einzelnen Cité oder durch die Kombination mehrerer Cités, deren moralische Prinzipien als kompatibel und gemeinschaftlich tragfähig ausgehandelt werden. Im Zentrum steht die politische Zielerreichung, die durch diese Verbindung als gesellschaftlich anerkennungswürdig legitimiert wird. Solche Kompromisse zeigen sich insbesondere dort, wo der Fokus nicht auf der Abgrenzung zwischen konkurrierenden Rechtfertigungen liegt, sondern etwa auf der Gestaltung eines funktionalen Einsatzes von KI, der verschiedene moralische Grundlagen integriert.

Empirische Analysen verschiedener politischer Kontroversen, wie der Finanzkrise (Gadinger & Yildiz, 2012), der Diskussion über die Verantwortung des UN-Sicherheitsrats während des Irak- sowie des Syrienkriegs (Niemann, 2019) oder der Enthüllungen Edward Snowdens (Smith Ochoa et al., 2021), machen dabei deutlich, wie insbesondere (aber nicht ausschließlich) politische Akteur:innen mithilfe von Rechtfertigungspraktiken um Legitimität ringen: „Importantly, it shows controversies (…) are characterised by a multitude of interpretations that strive for legitimacy via justificatory practices expressed through test formats in (non-)institutional contexts" (Smith Ochoa et al., 2021, S. 230). Dabei können im Laufe der Zeit und zwischen verschiedenen Akteur:innen nicht nur die Rechtfertigungspraktiken, also die Bezugnahme auf Cités, variieren und sich wandeln, sondern ebenso die Form der Kritik und das

Format der Prüfungen, das zur Herbeiführung legitimer Entscheidungen führen soll. Vor diesem Hintergrund wird somit die Kontingenz dieser Debatten deutlich (Smith Ochoa et al., 2021). Exemplarisch zeigt sich dies in den Auseinandersetzungen um den Einsatz von KI-Systemen in der öffentlichen Verwaltung. Während einige Akteur:innen deren Einführung zunächst mit Verweis auf Effizienz, Standardisierung und Objektivität rechtfertigen, wächst mit zunehmender Erfahrung häufig auch das Bewusstsein für potenziell diskriminierende Effekte und mangelnde Transparenz (Parycek et al., 2024). Damit treten Rechtfertigungen im Sinne der Cité der Industrie in ein Spannungsverhältnis zu Kritik, die sich auf Prinzipien der Cité des Staatsbürgers stützt. Diese Kritik kann je nach Kontext und Perspektive unterschiedlich stark ausfallen und von anwendungsspezifischen Einwänden bis hin zur grundsätzlichen, existenziellen Infragestellung des Einsatzes solcher Systeme reichen.

5.2.1. Kritik als konstitutiver Bestandteil der Aushandlung legitimer Entscheidugnen

Aufbauend auf den Annahmen der SdK sowie der Konzeptualisierung einer konfliktiven Politisierung wird Kritik einem agonistischen Verständnis folgend als konstitutiver Bestandteil legitimer Einigungsprozesse und demokratischer Gemeinschaften verstanden (Mouffe, 2013, 2015/2005; Straßenberger, 2020, 2021). Denn ohne Kritik befinden sich die Politik ebenso wie die Gesellschaft in einer scheinbar endlosen Schleife der Selbstbestätigung. Dabei kann Kritik verstanden werden als „Beitrag zur Konstitution der genannten Welten durch ihre Infragestellung“ (Bogusz, 2010, S. 56). Ohne eine solche Infragestellung würden der von Institutionen vorgegebene Handlungsrahmen und die damit verbundene Ordnung unreflektiert übernommen (Boltanski, 2008). Die Funktion der Kritik ist es daher, auf Konflikte und Probleme aufmerksam zu machen und im Zuge dessen dazu beizutragen, politische Entscheidungen und Rechtfertigungen daran anzupassen. In diesem Sinne dient beispielsweise die umfassende und beharrliche Technikkritik im Rahmen der STS (siehe Kapitel 6) als Antrieb eines gesellschaftlichen, technischen und politischen Wandels. Dabei ist ein stetes Wechselspiel zwischen Kritik und Rechtfertigung zu beobachten: „we can observe the dynamic tension between critique *as disturbance* to the centrality of innovation in contemporary governing institutions, and its *systemic reintegration* as a justificatory logic“ (Doezema & Frahm, 2023, S. 8; Herv. i. O.).

Kritik trägt gemäß den Annahmen der SdK in Situationen der Unsicherheit und Überforderung zur Komplexitätsreduktion bei, da Kritik stets unter Bezugnahme auf (mindestens) eine Cité geäußert wird, die als Koordinationsprinzip dient (Bogusz, 2010). Sie schließt somit im Sinne der Konfliktdimension an Politisierung an, indem sie zum einen Sachverhalte in den Raum des Politischen transportieren kann (Zürn, 2013) und zum anderen durch die Infragestellung der bestehenden Ordnung Antrieb der Rechtfertigung ist. Oder anders ausgedrückt: „Ohne Kritik

hat die Rechtfertigung keinen Sinn“ (Boltanski & Chiapello, 2001, S. 468). Gleichzeitig ist die Kritik selbst ebenso rechtfertigungspflichtig wie die Handlungen ihrer Adressat:innen, die sie beanstandet (Boltanski & Thévenot, 1991/2018; Niemann, 2019). Ebenso wie die Legitimität von Akteur:innen, einer Institution oder einer politischen Entscheidung kein Selbstläufer ist, der einmal erworben ohne weiteres Zutun bestehen bleibt, muss demnach auch die Kritik immer wieder erneuert und gerechtfertigt werden (Boltanski & Chiapello, 2001; Forst, 2015).

Dabei lassen sich grundsätzlich mit einer reformistischen und einer radikalen Form zwei Arten der Kritik unterscheiden (Boltanski, 2011; Boltanski & Chiapello, 2003; Forst, 2015). *Reformistische Kritik* zielt darauf ab, bestehende gesellschaftliche Ordnungen oder Entscheidungen nicht grundsätzlich zu verwerfen, sondern im Sinne einer Korrektur weiterzuentwickeln, weshalb sie auch als korrektive oder immanente Kritik bezeichnet wird (Böker, 2021). Sie hinterfragt so beispielsweise, ob eingelebte Sitten oder Institutionen ihren Zweck erfüllen, und zeigt Veränderungsbedarfe auf (Gadinger & Yildiz, 2012). So können etwa ethische Richtlinien, die auf gewisse Risiken des KI-Einsatzes hinweisen und Lösungsvorschläge zur Eindämmung selbiger vorlegen, als reformistische Kritik betrachtet werden, die auf eine Verbesserung der gegenwärtigen KI-Anwendung abzielt. *Radikale Kritik* ist hingegen revolutionistisch angelegt, indem sie den Zweck von Sitten, Institutionen oder bestehenden Praktiken gänzlich hinterfragt und zur Debatte stellt, ob dieser Zweck allgemein gerechtfertigt ist (Boltanski, 2008, 2010; Boltanski & Chiapello, 2003; Forst, 2015; Gadinger & Yildiz, 2012). Demnach werden weniger bestimmte Umgangsformen oder Handlungsoptionen kritisiert oder vorgeschlagen als vielmehr die Sinnhaftigkeit und Angemessenheit der von ihren (politischen) Gegner:innen praktizierten oder vorgeschlagenen Standpunkte gänzlich angezweifelt (Gadinger & Yildiz, 2012). Nicht selten stammt diese Kritik aus den Reihen marginalisierter Gruppen, die über persönliches Leid und Ungerechtigkeit klagen (Boltanski, 2011; Smith Ochoa et al., 2021). Im Kontext der Einführung von KI-Systemen kann radikale Kritik insbesondere von systematisch benachteiligten Gruppen geäußert werden, die die Anwendung solcher Technologien grundsätzlich infrage stellen, wie etwa im britischen Fall des Benotungsalgorithmus, bei dem massive Proteste schließlich zur Rücknahme des Systems führten (Smith, 2020).

5.2.2. Drei Prüfungssituationen der Aushandlung legitimer Entscheidungen

Boltanski (2011) unterscheidet in „*On Critique: A Sociology of Emancipation*“ drei Formen der Prüfung, die jeweils unterschiedliche Mechanismen zur Bewältigung von Unsicherheit darstellen. Sie dienen dazu, in sozialen Aushandlungsprozessen auf der einen Seite eine Einigung herzustellen und auf der anderen Seite bestehende soziale Ordnungen zu stabilisieren. Dabei

handelt es sich um Wahrheits-, Realitäts- und Existenzprüfungen, die sich sowohl in ihrer Funktion als auch in ihrem Verhältnis zur Kritik voneinander abgrenzen lassen (siehe Tabelle 1).

Tabelle 1. Übersicht der drei Prüfungsformen der Aushandlung nach Boltanski (2011)

	Wahrheitsprüfung	**Realitätsprüfung**	**Existenzprüfung**
Funktion	Bestätigung & Stabilisierung bestehender Ordnungen	Überprüfung der Einlösung normativer Ansprüche in der sozialen Realität	Fundamentale Infragestellung bestehender Ordnungen
Bezug zur Unsicherheit	Reduktion von Unsicherheit durch Beharren auf Stabilität und Kohärenz	Bearbeitung von Unsicherheit durch Aushandlung und Anpassung	Sichtbarmachung und Zuspitzung von Unsicherheit als Ausgangspunkt radikaler Kritik
Bezug zur Kritik	Kritik wird ausgeschlossen; Ordnung erscheint alternativlos	Reformistische Kritik zielt auf Verbesserung oder Anpassung	Radikale Kritik richtet sich gegen die Legitimität der Ordnung selbst
Typische Akteur:innen	Regierungen	Opposition, Expert:innen	Zivilgesellschaft, marginalisierte Gruppen
Prüfungslogik	Tautologisch, selbstreferenziell, normstabilisierend	Differenzierende Bewertung des „Solls“ und „Seins“	Emotional, subjektiv, moralisch radikal
Angestrebtes Ergebnis	Anerkennung und Festigung der bestehenden Legitimitätsordnung	Anpassung, Reform oder Kompromissbildung	Transformation oder Delegitimierung bestehender Ordnungen
Exemplarischer KI-Bezug	Rhetorik politischer Kontrolle und Sicherheit der KI	Debatten über Regulierungsanforderungen an KI	Fundamentaler Widerstand gegen den KI-Einsatz

Wahrheitsprüfungen zielen auf die Bestätigung eines vermeintlich unumstößlichen „präetablierten Zustand[s]“ (Boltanski, 2008, S. 138) ab. Sie verzichten auf Kritik als emanzipatorische Kraft und stützen sich stattdessen auf Tautologien und Selbstbestätigungen (Boltanski, 2010, 2011; Gadinger & Yildiz, 2012): „Ihre einzige Rolle besteht darin *zu zeigen, daß es eine Norm gibt*“ (Boltanski, 2010, S. 155; Herv. i. O.). Damit vermitteln sie einen Eindruck von Notwendigkeit und innerer Kohärenz, der durch ständige Wiederholung überzeugend wirken soll (Boltanski, 2010). Diese Form der Prüfung ist damit kein Mittel der Veränderung, sondern der Stabilisierung, wobei Kritik systematisch ausgeschlossen wird. Gelingt die Wahrheitsprüfung, kommt es zu einer Anerkennung der konstruierten Wirklichkeit sowie der ihr zugrunde liegenden Cités, ohne dass diese weiter geprüft werden müssten. Somit würde die Komplexitätsreduktion durch die Sicherung der Stabilität der sozialen Ordnung erreicht.

Trotz dieser Kohärenz behält Boltanski (2008) die Bezeichnung einer *Prüfung* bei, da auch Wahrheitsprüfungen scheitern können. Dies gilt insbesondere dann, wenn die symbolische Ordnung, die sie absichern sollen, durch unerwartete Entwicklungen oder Ereignisse infrage

gestellt wird. So können etwa technologische Neuerungen wie KI-Systeme, die den ihnen zugeschriebenen Erwartungen möglicherweise nicht gerecht werden, Unsicherheit erzeugen (Boltanski, 2010). Wahrheitsprüfungen können somit an der sozialen Realität scheitern, wenn beispielsweise unintendierte Effekte des KI-Einsatzes zutage gefördert werden und somit den Anschein der Stabilität erschüttern.

In politischen Debatten bedienen sich insbesondere Regierungsparteien häufig der Wahrheitsprüfung, indem sie auf Metaphern zurückgreifen, die Sicherheit und Stabilität suggerieren. Auf diese Weise wird der Eindruck vermittelt, sämtliche Entwicklungen seien kontrollierbar. Es handelt sich um ein rhetorisches Mittel, das dem Machterhalt und der Absicherung bestehender Ordnungen dient (Smith Ochoa et al., 2021). Im Zentrum dieser Praktiken steht die Legitimation des Status quo, nicht jedoch dessen kritische Hinterfragung oder die Eröffnung alternativer Handlungsoptionen (Gadinger & Yildiz, 2012). Vor diesem Hintergrund kann der Einsatz neuer KI-Technologien als ein Moment der Verunsicherung verstanden werden, der etablierte Ordnungsvorstellungen infrage stellt. Die potenziell destabilisierenden Effekte solcher Technologien werden im Rahmen von Wahrheitsprüfungen rhetorisch neutralisiert. Indem Unsicherheit ausgeblendet und Kritik unterbunden wird, soll die bestehende Legitimitätsordnung aufrechterhalten werden. Genau hierin liegt jedoch eine zentrale Problematik: Die zugrunde liegende Cité bleibt unangetastet und reproduziert sich in einer zirkulären Struktur fort, die auf Selbstbestätigung und die Ausblendung von Alternativen angewiesen ist (Boltanski, 2008, 2010).

Realitätsprüfungen stellen im Vergleich zur Wahrheitsprüfung eine grundlegend andere Form des Umgangs mit Unsicherheit dar. Im Mittelpunkt steht nicht die Bestätigung eines bestehenden Ordnungsmodells, sondern die Prüfung, ob die normativen Ansprüche einer herangezogenen Cité in der sozialen Realität eingelöst werden können. Während Wahrheitsprüfungen das „Wünschenswerte so [entfalten], als sei es gegeben“ (Boltanski, 2010, S. 159), zielen Realitätsprüfungen auf eine Differenzierung „zwischen Sollen und Sein, zwischen Werturteil und Tatsachenurteil“ (ebd.). In diesem Spannungsverhältnis zwischen Ist-Zustand und normativem Soll-Zustand eröffnen sie einen Raum für Kritik, Reflexion und potenzielle Neujustierungen sozialer Ordnungen (Bogusz, 2010; Boltanski, 2010, 2011; Forst & Günther, 2011; Smith Ochoa et al., 2021).

Realitätsprüfungen initiieren häufig Aushandlungsprozesse über die Gültigkeit oder konkrete Umsetzung normativer Maßstäbe. Dabei richtet sich die Kritik nicht gegen die Existenz der jeweiligen Cité selbst, sondern gegen ihre unzureichende Realisierung oder gegen konkurrierende Rechtfertigungsordnungen (Boltanski, 2011). Im Zentrum steht die reformistische Kritik,

die bestehende Normen an veränderte gesellschaftliche Kontexte anpassen will, etwa durch die Formulierung alternativer Regeln oder politischer Maßnahmen zur Verbesserung der Ausgangssituation (Boltanski, 2008, 2011; Gadinger & Yildiz, 2012). So manifestieren sich Realitätsprüfungen empirisch betrachtet typischerweise in Wortbeiträgen der Opposition oder in publizistischen Kommentaren (Smith Ochoa et al., 2021). Je nach Ausrichtung kann eine Realitätsprüfung stabilisierend wirken, indem sie bestehende Ordnungen bekräftigt, oder transformativ, wenn sie zu deren Korrektur beiträgt (Boltanski, 2010). Entsprechend bezeichnet Gadinger (2016) Realitätsprüfungen als Kern der Legitimationspraxis: „The typical legitimacy struggle is a 'reality test,' in which different interpretations of reality, objects, environments, and devices must be incorporated in practices of justification and critique" (S. 198). Spannungen zwischen normativen Ansprüchen und realen Gegebenheiten sowie zwischen Kritik und Rechtfertigungspraktiken werden in diesem Zusammenhang häufig durch das Aushandeln von Kompromissen bearbeitet (Boltanski, 2010).

Boltanski und Thévenot (1991/2018) veranschaulichen diese Dynamik am Beispiel familiärer Übergangssituationen, wie etwa Hochzeiten oder Todesfälle, in denen die häusliche Cité unter Druck gerät. Solche Ereignisse können etablierte Rangordnungen herausfordern und erfordern entweder deren Bestätigung oder eine Neuverhandlung. Auch hier bleibt die Gültigkeit der Cité meist unbestritten, während sich die Auseinandersetzung auf deren konkrete Ausgestaltung fokussiert. Vor diesem Hintergrund können Prüfungen als „Bewährungsproben" verstanden werden, die eine Transformation ermöglichen, ohne notwendigerweise revolutionären Charakter zu haben (Boltanski, 2008). So erscheint es naheliegend, dass sich Realitätsprüfungen im Kontext von KI-Debatten typischerweise in der Frage manifestieren, *wie* KI-Systeme implementiert werden sollten, und weniger darin, *ob* dies grundsätzlich geschehen sollte. Die zugrunde liegende normative Ordnung wird nicht infrage gestellt, vielmehr richtet sich die Debatte auf die praktische Umsetzung, also auf die Frage, wie moralische Prinzipien bei der Anwendung technologischer Systeme berücksichtigt werden können.

Existenzprüfungen unterscheiden sich hingegen grundlegend von den zuvor beschriebenen Prüfungsformen. Im Gegensatz zur stärker rational argumentierenden Realitätsprüfung oder zur normstabilisierenden Wahrheitsprüfung sind sie durch ein hohes Maß an Subjektivität, Emotionalität und radikaler Kritik gekennzeichnet. Im Zentrum steht hier nicht die Anpassung bestehender Ordnungssysteme, sondern die grundlegende Infragestellung ihrer Legitimität. Existenzprüfungen artikulieren Ansprüche auf Gerechtigkeit und Anerkennung, die auf persönlichen Erfahrungen gründen und auf die öffentliche Sichtbarmachung bislang marginalisierter

Perspektiven zielen (Boltanski, 2011). Dabei werden die Kontingenz bestehender Verhältnisse offengelegt und alternative Handlungsoptionen eingefordert (Boltanski, 2008).

Ein exemplarisches Fallbeispiel bietet Boltanski (2010) mit der Situation einer Mutter, die ihrem querschnittsgelähmten Sohn auf dessen Wunsch hin Sterbehilfe leistet und daraufhin strafrechtlich verfolgt werden soll. Die normative Ordnung, dass Tötung ein Verbrechen ist, wird durch diese existenzielle Konstellation selbst zum Gegenstand der Prüfung, indem eine geltende Definition des „Verbrechens" hinterfragt und neu ausgehandelt werden muss. Die Darstellung „Wenn diese Frau eine Verbrecherin ist, dann ist jede liebende Mutter eine Verbrecherin" (Boltanski, 2010, S. 164) veranschaulicht die moralische Sprengkraft, die von Existenzprüfungen ausgeht.

Aufgrund ihrer radikalen Infragestellung etablierter Legitimitätsordnungen finden Existenzprüfungen selten Eingang in formalisierte politische Verfahren. Stattdessen werden sie vorrangig im Kontext zivilgesellschaftlicher Bewegungen, künstlerischer Praktiken oder über soziale Netzwerke artikuliert, das heißt in Arenen, die institutionelle Routinen unterlaufen (Gadinger, 2016; Smith Ochoa et al., 2021). Dort richten sich diese Prüfungen nicht nur gegen einzelne politische Entscheidungen, sondern gegen die Legitimität der ihnen zugrunde liegenden Institutionen und Ordnungen: „mainly articulated outside the parliamentary routines by artists, intellectuals, or ordinary actors in questioning the very legitimacy and existence of a particular institution or instrument therein" (Gadinger, 2016, S. 201). Die emotionale Aufladung radikaler Kritik kann sich sowohl in utopischen als auch in dystopischen Zukunftsvorstellungen niederschlagen. Personen, Ereignisse oder eben auch Technologien wie KI erscheinen hier entweder als Heilsversprechen oder als Bedrohung, wobei sie je nach Perspektive als Lösung für strukturelle Ungleichheiten oder als Verstärker ebenjener Ungerechtigkeiten inszeniert werden (Smith Ochoa et al., 2021). Forderungen nach einem vollständigen Verbot diskriminierender KI-Systeme ließen sich insofern als Ausdruck einer existenziellen Prüfung deuten, die nicht auf Reform, sondern auf eine grundsätzliche Infragestellung und fundamentalen Wandel zielt.

5.3. Konzeptioneller Rahmen der Analyse der Legitimation der politischen Entscheidung über den Einsatz künstlicher Intelligenz

An dieser Stelle lässt sich ein Zwischenfazit ziehen, um den konzeptionellen Rahmen zur Analyse der Legitimation politischer Entscheidungen über den Einsatz von KI im Bildungssystem zusammenzufassen (siehe Abbildung 1). Ausgangspunkt der Arbeit ist die forschungsleitende Frage, wie politische Akteur:innen Entscheidungen über den KI-Einsatz legitimieren, obwohl erhebliche Unsicherheiten über deren potenziell weitreichende Folgen bestehen. Es wurde aufgezeigt, dass Politisierung dabei eine notwendige Bedingung legitimer Entscheidungsfindung

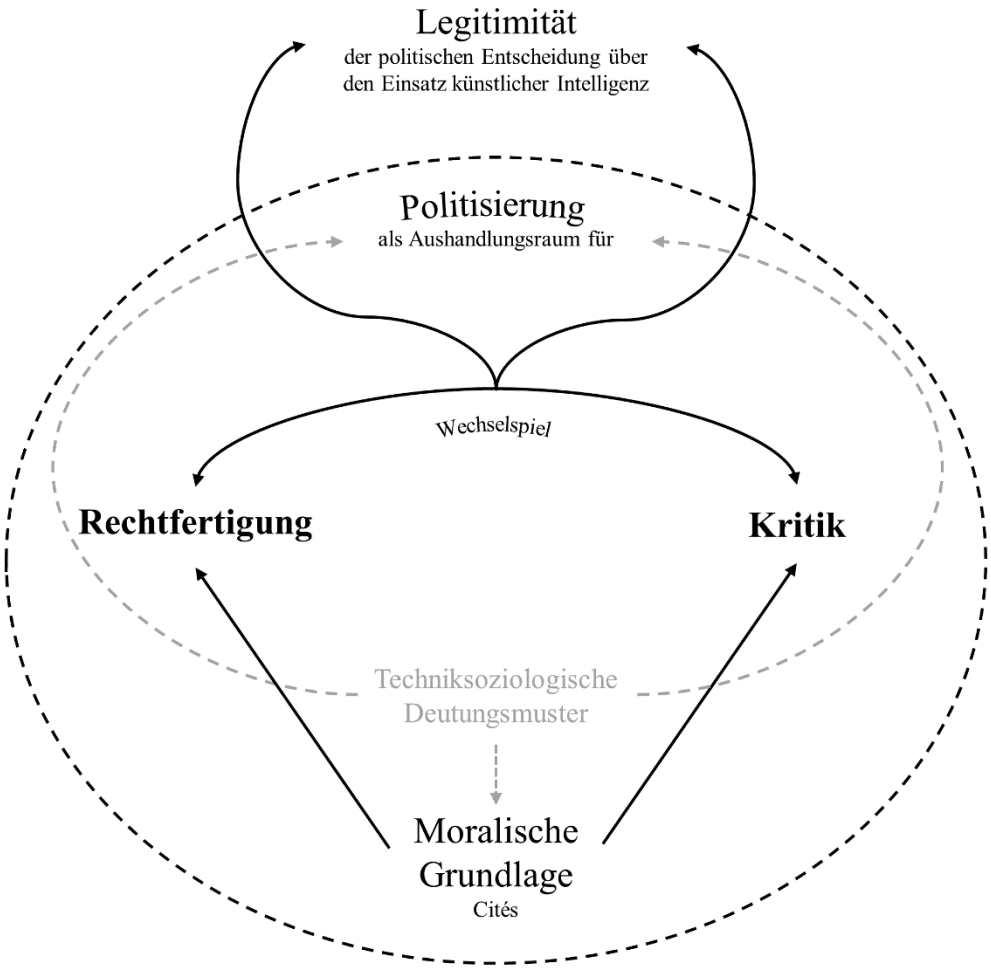
darstellt. Sie eröffnet einen kontingenten und konflikthaften Aushandlungsraum, in welchen unterschiedliche Perspektiven eingebracht und bewertet werden können. Damit wird sichergestellt, dass der Einsatz von KI als politisch gestaltbares Handlungsfeld anerkannt wird. Die Anerkennung dieser Kontingenz und Gestaltbarkeit erzeugt zugleich einen erhöhten Rechtfertigungsdruck, der politische Akteur:innen dazu verpflichtet, ihre Entscheidungen öffentlich zu begründen. Die Einführung von KI im Bildungswesen wird in diesem Zusammenhang als kritischer Moment betrachtet, der die bestehende Ordnung herausfordert und neue Handlungsspielräume schafft. Diese erfordern eine Aushandlung moralischer Grundlagen und möglicher Kompromisse, um durch den KI-Einsatz herausgeforderte Normen zu bestätigen oder neu zu bestimmen.

An diesem Punkt setzt das pragmatische Legitimitätsverständnis der SdK an. Es begreift legitime Entscheidungsfindung als ein Wechselspiel zwischen Rechtfertigung und Kritik, das sich insbesondere in politisierten Situationen normativer Unsicherheit entfaltet. Während die Rechtfertigung der moralischen Begründung des eigenen Standpunkts dient, ermöglicht Kritik, alternative Handlungsspielräume aufzuzeigen und den Rechtfertigungsdruck aufrechtzuerhalten. Zur moralischen Fundierung ihrer Positionen greifen Akteur:innen auf Cités zurück. Diese dienen als Ordnungsschema zur Bewertung von Handlungen und Entscheidungen und bieten Orientierung im Umgang mit Komplexität und Ungewissheit. Sie strukturieren politische Debatten, indem sie auf gesellschaftlich geteilte Konventionen rekurrieren, die als nachvollziehbare Bezugspunkte für Kritik und Rechtfertigung dienen. Der Rückgriff auf Cités ermöglicht eine koordinierte Bezugnahme in der politischen Debatte. Akteur:innen können entweder eine Einigung darüber anstreben, welche Cité in der konkreten Situation als normativ angemessen gilt, oder einen Kompromiss aushandeln, der verschiedene moralische Orientierungen kombiniert. Solche Kompromisse beruhen auf der wechselseitigen Anerkennung unterschiedlicher Geltungsansprüche, die als kompatibel und gemeinschaftlich tragfähig ausgehandelt werden. In beiden Fällen trägt das Zusammenspiel von Kritik und Rechtfertigung zur Herstellung legitimer politischer Entscheidungen bei. Die Analyse dieses moralisch fundierten Wechselspiels bildet den Kern der empirischen Untersuchung der Arbeit.

Wie die Diskussion des Politisierungspotenzials von KI angedeutet hat (siehe Abschnitt 4.2.3), ist dabei davon auszugehen, dass politische Positionierungen durch Vorstellungen über die technische Gestalt und Wirkung von KI beeinflusst werden. Diese Vorstellungen können Einfluss darauf nehmen, welche moralischen Grundlagen als legitim erscheinen. Zum Ausdruck kommen sie in techniksoziologischen Deutungsmustern, die sowohl die moralische Rechtfertigung des politischen Umgangs mit KI als auch die wahrgenommene Kontingenz politischer

Gestaltbarkeit im Sinne der Politisierung beeinflussen können. Vor diesem Hintergrund wird im nachfolgenden Kapitel auf Grundlage des aktuellen Forschungsstandes in den Blick genommen, inwiefern technikdeterministische, sozialkonstruktivistische oder technopragmatische Deutungen die moralische Grundlage von Rechtfertigung und Kritik im Kontext des KI-Einsatzes strukturieren können.

Abbildung 1. Konzeptioneller Rahmen zur Analyse legitimer politischer Entscheidungsfindung des Einsatzes künstlicher Intelligenz



6. Techniksoziologische Deutungsmuster als Legitimationsgrundlage moderner Technologien

Vorherrschende Deutungen von Technik und ihrer gesellschaftlichen Wirkung prägen maßgeblich die öffentliche Kritik und Rechtfertigung einer Technologie und bilden damit die Grundlage gesellschaftlicher Realität und auch politischer Regulierung (Bareis & Katzenbach, 2022; Borup et al., 2006; Brause et al., 2025; Jasanoff, 1987, 2015; Natale & Ballatore, 2020; Selin, 2007; van Lente & Rip, 1998; van Lente et al., 2013; Zerbe, 2007). Eine historische Betrachtung dieser Zusammenhänge zeigt, dass diese Technikdeutungen zur Legitimationsgrundlage politischer Entscheidungen werden (Bogner, 2021b; Rahm, 2023a, 2023b). Vor diesem Hintergrund

lassen sich die Muster der Kritik und Rechtfertigung moderner Technologien als Ausdruck vorherrschender techniksoziologischer Grundannahmen über das Verhältnis von Technik und Gesellschaft konzipieren. Diese Grundannahmen prägen somit die Legitimationsgrundlage technischer Innovationen. Ziel dieses Kapitels ist es daher, aufzuzeigen, inwiefern techniksoziologische Paradigmen politische Aushandlungsprozesse leiten und somit erstens die Rechtfertigung und Kritik technischer Innovationen, zweitens den Grad ihrer Politisierung und damit drittens die Legitimation technologiepolitischer Entscheidungen beeinflussen.

Die techniksoziologische Debatte ist lange Zeit durch ein Spannungsverhältnis zwischen technikdeterministischen und sozialkonstruktivistischen Paradigmen geprägt und widmet sich damit der Frage, ob Technik Gesellschaft beeinflusst oder Gesellschaft die technische Entwicklung (Degele, 2002; Häußling, 2019; Passoth, 2008; Rammert, 2016; Schulz-Schaeffer, 2008; Weyer, 2008). Während technikdeterministische Perspektiven Technik als treibende Kraft gesellschaftlicher Veränderung begreifen, betonen sozialkonstruktivistische Ansätze die gesellschaftliche Gestaltung technischer Entwicklungen. In der wissenschaftlichen Debatte wird dieses Spannungsverhältnis als Kontinuum verstanden, entlang dessen sich unterschiedliche techniksoziologische Positionen verorten lassen (Dafoe, 2015; Marx & Smith, 1994). So versuchen etwa jüngere Ansätze des Technopragmatismus (Rammert, 2016), diesen Dualismus zu überwinden, indem sie ein „pragmatisches Verständnis von Technik [vertreten], das die Ko-Konstitution von Technik und Sozialem betont“ (Anicker, 2023, S. 140). Gleichzeitig bleiben sowohl technikdeterministische als auch sozialkonstruktivistische Technikdeutungen beispielsweise in der medialen KI-Debatte weiterhin präsent (Winkel, 2025).

In der wissenschaftlichen Debatte lässt sich eine historische Lernbewegung erkennen, in deren Verlauf deterministische Technikdeutungen zunächst durch sozialkonstruktivistische und später durch technopragmatische Ansätze relativiert und weitgehend abgelöst wurden (Passoth, 2008). Diese Verschiebung verweist auf einen Erkenntnisgewinn im Umgang mit Technik, der aus zunehmenden gesellschaftlichen Erfahrungen mit ihren Wirkungen hervorgegangen ist. Auch wenn diese Perspektiven keiner strikt trennscharfen Chronologie folgen und in der gegenwärtigen technologiepolitischen wie techniksoziologischen Debatte weiterhin nebeneinander bestehen, zeigt sich historisch betrachtet, dass bestimmte Technikdeutungen im Verlauf technischer Entwicklungen an Plausibilität verlieren und schrittweise durch andere Perspektiven ersetzt werden. Diese Dynamik lässt sich sowohl im Wandel der Technologiepolitik seit der Nachkriegszeit (Weyer, 2008) als auch in typischen Innovationszyklen beobachten, in denen Phasen hoher Erwartungen regelmäßig in Ernüchterung und Neubewertung münden (Geels & Verhees, 2011; Haenlein & Kaplan, 2019; Hirsch-Kreinsen, 2024). Eine historische

Perspektive eröffnet daher wertvolle Analogien für das Verständnis gegenwärtiger Formen der Kritik, Rechtfertigung und Politisierung von KI.

Im Folgenden werden die Zusammenhänge zwischen techniksoziologischen Deutungsmustern und technologiepolitischen Entscheidungen systematisch entfaltet. Die Abschnitte folgen dabei einer einheitlichen Struktur: Zunächst wird das jeweilige techniksoziologische Deutungsmuster historisch eingeordnet und es wird aufgezeigt, mit welchen Rechtfertigungsformen es traditionell verknüpft werden kann. Darauf aufbauend werden theoretische und empirische Befunde zur politischen KI-Debatte herangezogen, um diese Deutungsmuster und die damit verbundenen Rechtfertigungen sowie die gegenwärtige techniksoziologische Kritik im aktuellen Kontext einzuordnen. Abschließend wird auf Grundlage historischer Erfahrungen herausgearbeitet, welche Konsequenzen sich aus den jeweiligen Deutungsmustern für die Politisierung moderner Technologien ergeben. Im Mittelpunkt dieser Rekonstruktion stehen dabei nicht die bereits dargelegten allgemeinen zugesprochenen Potenziale und Risiken (siehe Abschnitte 2.2 und 2.3) sowie die daran anschließende Frage, *wie* KI-Systeme „gut" oder „gerecht" gestaltet werden können (siehe Abschnitt 2.4), sondern *ob* und in welchem gesellschaftlichen Rahmen ihr Einsatz legitim erscheint. Dabei werden Bezüge zu den Cités (siehe Abschnitt 5.1) entlang von vier techniksoziologischen Perspektiven herausgearbeitet (siehe Tabelle 2): (a) das Fortschrittsversprechen des optimistischen Determinismus, (b) der Kontrollverlust des pessimistischen Determinismus, (c) die Demokratisierung des Sozialkonstruktivismus und (d) die moralische Fundierung der Technikgestaltung des kritischen Pragmatismus.

Die Ausführungen stützen sich auf techniksoziologische Literatur sowie auf empirische Untersuchungen der politischen KI-Debatte. Im Rahmen der SdK liegen bislang vor allem theoretische Auseinandersetzungen zu den moralischen Grundlagen von Metrisierungsprozessen und KI-Anwendungen vor (Diaz-Bone & Horvath, 2021; Diaz-Bone et al., 2020; Diaz-Bone & Schrör, 2022; Frei et al., 2023; Horvath et al., 2025; Kappler, 2025), während empirische Studien zur politischen KI-Debatte begrenzt sind und sich überwiegend auf Rechtfertigungen wirtschaftlicher Akteur:innen konzentrieren (Kajava et al., 2025; Nachtwey & Seidl, 2023). Die vorliegende Arbeit zielt darauf, diese Leerstelle für die bildungspolitische Debatte zu adressieren, und greift an dieser Stelle ergänzend auf breiter angelegte Studien zurück, die dominante Framings und Argumentationslogiken im KI-Kontext identifizieren. Diese Befunde dienen als analytische Brücke, um die Plausibilität unterschiedlicher moralischer Rechtfertigungsgrundlagen theoretisch zu fundieren und empirisch abzusichern. Auf dieser Grundlage werden die zentralen Desiderate für die vorliegende Arbeit abgeleitet (siehe Kapitel 7).

Tabelle 2. Techniksoziologische Deutungsmuster, ihre Kritik- und Rechtfertigungsmuster und ihre Konsequenzen für die Politisierung von Technik

Techniksoziologische Deutungsmuster	**Historische Rechtfertigungsperspektive(n)**	**Politisierung**
Optimistischer Determinismus (i) Annahme: Eigenlogik der Technik als treibende Kraft gesellschaftlichen Fortschritts (ii) Ausgangspunkt: Hohe Technikgläubigkeit & Fortschrittsoptimismus/-utopie (iii) Konsequenz: Konfliktfreiheit, Technologische Entwicklung naturwüchsig; politische Gestaltung kaum erforderlich	**Rechtfertigung** - Industrie: Effizienzgewinne; Planungssicherheit; Innovationsstreben - Markt: Wettbewerbsbestrebungen - Solution: Problemlösungsfähigkeit - Inspiration: individuelle Freiheitsgewinne	**Geringe Politisierung** - Kein Kontingenzbewusstsein politischer Gestaltbarkeit - Konfliktfreiheit aufgrund Annahme inhärenter Technik-Legitimität
Pessimistischer Determinismus (i) Annahme: Eigendynamik der Technik führt zu Kontrollverlust & Zwängen (ii) Ausgangspunkt: Wachsendes Bewusstsein über Risiken & unintendierte Folgen durch technische Fehlschläge stellt Legitimität des Technikeinsatzes infrage (iii) Konsequenz: Technik- & Technokratiekritik; wachsende Polarisierung zw. Technikoptimismus und -pessimismus	**Kritik** - Staatsbürger: Kontroll- & Autonomieverlust des Menschen durch Eigendynamik & Zwänge der Technik; Technokratie widerspricht Ideal mündiger Bürger:innen	**Reaktive Politisierung** - Wachsende öffentliche Auseinandersetzung über Technikfolgen & -risiken infolge negativer Erfahrungen
Sozialkonstruktivismus (i) Annahme: Soziale Gestaltbarkeit & Interpretationsbedürftigkeit der Technik (ii) Ausgangspunkt: Anerkennung politischer Qualität von Technik; empirische Erkenntnisse des gesellschaftlichen Einflusses auf Technikentwicklung; Kritik an technokratischer Steuerung (iii) Konsequenzen: Forderung aktiver gesellschaftlicher Beteiligung	**Rechtfertigung** - Staatsbürger: Demokratisierung der Technik durch Inklusion & Partizipation **Kritik** - Staatsbürger: Gefährdung demokratischer Grundrechte - Ökologie: Vernachlässigung gesundheitlicher & ökologischer Risiken	**Beabsichtigte Politisierung** - Konstruktivistische Perspektive fordert aktive Gestaltung von Technik
Kritischer Pragmatismus (i) Annahme: Wechselseitige Beeinflussung von Technik & Gesellschaft; Technik als Ergebnis dynamischer Ko-Evolution (ii) Ausgangspunkt: Kritik rein partizipativer Ansätze; Einbettung technischer Innovation in Netzwerkpolitik (iii) Konsequenzen: Forderung sozialverantwortlicher Technikgestaltung	**Rechtfertigung** - Staatsbürger: Gemeinwohlorientierung; Chancengleichheit; Partizipation - Ökologie: Nachhaltigkeit **Kritik** - Staatsbürger: Technik inkorporiert hegemoniale Machtverhältnisse; unsachgemäße Komplexitätsreduktion sozialer Probleme	**Gefordert Politisierung** - Kritischer Pragmatismus erfordert Kritik als Grundlage der Legitimation der Technik

6.1. Das unausweichliche Fortschrittsversprechen der Technik: Die Perspektive des optimistischen Determinismus

Eine historische Betrachtung zeigt, dass gerade in frühen Phasen technischer Innovationen diese in der öffentlichen Wahrnehmung häufig mit hohen, mitunter auch utopischen Erwartungen verknüpft sind (Brown, 2003; van Lente et al., 2013). So begegneten etwa Politik und Öffentlichkeit der Atomkraft zu Beginn mit einer „unerschütterlichen Technikgläubigkeit“ (Schubert, 1991, S. 123) und einer ausgeprägten Faszination für ihre technische Leistungsfähigkeit (Keck, 1984a, 1984b). Diese Begeisterung gründet auf einem technikdeterministischen Paradigma, das mit optimistischen Fortschrittsversprechen verknüpft ist. Es beruht auf der Annahme, dass technische Innovationen einer eigenen Dynamik folgen, die sich weitgehend unabhängig von sozialen Einflüssen entfaltet. Technik erscheint dabei als treibende Kraft gesellschaftlichen Wandels, an die sich soziale Strukturen anzupassen haben (Degele, 2002; Ogburn, 1969). Vor allem in der Nachkriegszeit galt Technik dabei sowohl in der techniksoziologischen Bewertung als auch in der politischen Debatte als Motor des Fortschritts und als zentrales Instrument zur Steigerung von Wohlstand und Effizienz (Dessauer, 1927; McLuhan, 1964; Veblen, 1912/1997).

Diese optimistisch-deterministische Perspektive, obwohl tief im letzten Jahrhundert verwurzelt, findet bis heute Widerhall in der öffentlichen und politischen Auseinandersetzung über KI (Ballatore & Natale, 2023; Campolo & Crawford, 2020; Rahm, 2023a). In nationalen KI-Strategien wird KI häufig als unvermeidlicher Einflussfaktor beschrieben, der Gesellschaft und Politik tiefgreifend verändern wird (Ulnicane et al., 2022). Sowohl in der deutschen als auch in der französischen, US-amerikanischen und chinesischen KI-Politik gilt der technologische Wandel als unausweichlich:

> [T]hey all establish AI as a given and massively disrupting technical development that will change society and politics fundamentally. In consequence, the necessity to adopt AI across all key sectors of society is portrayed as taken for granted and inevitable. (Bareis & Katzenbach, 2022, S. 875)

Neben politischen Akteur:innen betonen insbesondere wirtschaftliche Akteur:innen, die sich selbst als Profiteur:innen der KI-Durchdringung verstehen, die transformative Kraft und scheinbare Unvermeidlichkeit von KI (Schopmans, 2022). Vertreter:innen dieser optimistischen Position verweisen dabei auf umfassende Fortschritts- und Wohlstandsversprechen (Brynjolfsson & McAfee, 2014), die im Kontext von KI in bislang noch nie dagewesenem Ausmaß wirksam werden sollen (Russell et al., 2015).

Diese technikdeterministische Grundannahme verband sich in der Nachkriegszeit mit einem optimistischen Fortschrittsglauben, der im Konzept des „*technological fix" (A. M. Weinberg, 1967)* seinen Ausdruck fand. Demnach lassen sich gesellschaftliche Probleme durch technologische Innovationen effizienter lösen als durch soziale oder politische Maßnahmen. Dieses Verständnis knüpft an technokratische Vorstellungen an, nach denen moderne Demokratien in hohem Maße auf die Expertise von Fachleuten angewiesen sind (Johnston, 2017). In der gegenwärtigen KI-Debatte lebt diese Idee fort in Form von Konzepten wie *Solutionismus* (Morozov, 2013) oder *Technochauvinismus* (Broussard, 2018; König & Wenzelburger, 2022). Technologische Entwicklungen, insbesondere KI-Anwendungen, werden dabei als bevorzugte oder gar einzig sinnvolle Antwort auf gesellschaftliche Herausforderungen interpretiert (Bartl, 2023; Nachtwey & Seidl, 2023; Sætra & Selinger, 2024). Die Idee des Solutionismus ist zugleich Ausdruck eines Ringens um epistemische Autorität bei der Definition effizienter Problemlösungsstrategien. Vor allem Technologieunternehmen beanspruchen diese Deutungshoheit, indem sie KI-Anwendungen als Werkzeuge darstellen, die gesellschaftliche Effizienz und politische Zielerreichung gleichermaßen fördern sollen, und so technologische Innovation mit Gemeinwohlorientierung verknüpfen (Obendiek & Seidl, 2023).

6.1.1. Rechtfertigung technologischer Innovation zwischen Effizienz, Wettbewerb und Allheilmittel

Diese Form der Technikeuphorie prägte in der historischen Betrachtung lange Zeit die politische Rahmung des technologischen Fortschritts, die auf verschiedene Cités verweist. Das Versprechen von Effizienz, Produktivität und Planungssicherheit durch Rationalisierung und Automatisierung (Rahm, 2023a; Veblen, 1912/1997; Volti & Croissant, 1988/2024) kann etwa als zentraler Bezug auf die Cité der Industrie verstanden werden. Gleichzeitig wirkten zu Zeiten des Kalten Krieges geopolitische Faktoren wie der Wettstreit zwischen Ost und West als Antrieb für technische Großprojekte mit nationalem Prestige und ökonomischer Stärke (Edler & Wittmann, 2024; Mai, 2011; Weyer, 2008), die auf marktorientierte Rechtfertigungen verweisen. Ergänzend spielte in der Nachkriegszeit ebenso die politische und gesellschaftliche Vorstellung individueller Befreiung durch technische Entlastung eine Rolle (Bogner, 2021b), die an die Cité der Inspiration anschließt.

Diese historischen Technikdeutungen eines optimistischen Determinismus bleiben auch weiterhin aktuell, wenngleich die Benennung konkreter Potenziale zunehmend diffus erscheint (Rahm, 2023a). Dabei wird KI in der öffentlichen Debatte sowohl als Allheilmittel, das gesellschaftliche und politische Probleme technologisch zu lösen verspricht, diskutiert als auch im Kontext globaler Anschlussfähigkeit und ökonomischen Wohlstands.

6.1.1.1. Der Aufschwung des Solutionismus: Künstliche Intelligenz gerechtfertigt als Allheilmittel

Der optimistische Determinismus der Nachkriegszeit spiegelt sich in der gegenwärtigen techniksoziologischen KI-Debatte in der Auseinandersetzung mit der Verbreitung des Solutionismus wider. Dieser bezeichnet eine ideologisch geprägte Denkweise, deren Entstehung und Durchsetzung sowohl durch den digitalen Wandel und den Aufstieg globaler Technologiekonzerne als auch durch den neoliberalen Umbau staatlicher Strukturen begünstigt wird, in dessen Zuge Marktmechanismen und Effizienzlogiken an Bedeutung gewinnen (Morozov, 2016). So verstärken sich Neoliberalismus und Solutionismus wechselseitig, indem insbesondere die im Silicon Valley dominante Technikdeutung KI als zentralen Weg gesellschaftlicher Problemlösung darstellt und damit grundlegende Annahmen neoliberaler Politik stützt, die soziale Konflikte entpolitisieren und ihre Bearbeitung in Marktmechanismen verlagern (Ferrari, 2020). Der gegenwärtige KI-Hype wird somit maßgeblich durch Technologie-Unternehmen angetrieben, die nahezu ausschließlich auf die positiven Wirkungen der Technologie vertrauen. Dabei wird die Profitabilität von KI mit einem Wohlstandsversprechen sowie der Vision eines Problemlösungsinstruments verknüpft (Eynon & Young, 2021; Horvath et al., 2025; Khanal et al., 2025; Nachtwey & Seidl, 2023; Sætra & Selinger, 2024; Williamson & Komljenovic, 2023). Dadurch verschmelzen ökonomische, politische und moralische Legitimationsansprüche zu einem einheitlichen Narrativ, das technologische Lösungen als im Sinne des Gemeinwohls handelnd legitimiert (Nachtwey & Seidl, 2023). Vertreter:innen einflussreicher Technologiekonzerne (die sogenannte neue Digital-Elite) propagieren die grenzenlosen Einsatzmöglichkeiten von KI (Burrell & Fourcade, 2021; Stilgoe, 2020; Williamson & Komljenovic, 2023). Ihre Versprechen reichen von schnellerer bis hin zu vermeintlich fairerer Entscheidungsfindung im Vergleich zu menschlichen Verfahren (Burrell & Fourcade, 2021) und zielen auf eine moralische Aufladung technologischer Entwicklung, indem sie Fortschritt mit sozialem Nutzen gleichsetzen, wobei KI zum Beispiel als Beitrag zu mehr Sicherheit oder Gerechtigkeit inszeniert wird (Brockmann et al., 2021; Obendiek & Seidl, 2023). Explizite Analysen der zugrunde liegenden Cités verweisen darauf, dass die Digital-Elite ihre Legitimation wesentlich auf solutionistische Rechtfertigungen stützt, das heißt auf die Vorstellung, dass technologische Innovation per se gesellschaftliche Probleme lösen könne (Nachtwey & Seidl, 2023). Dieses Rechtfertigungsmuster wird von wirtschaftlichen Akteur:innen durch ökonomische Wettbewerbsnarrative und Effizienzversprechen stabilisiert, die den Glauben an technische Rationalität und Steuerbarkeit weiter stärken und den Cités der Industrie und des Marktes zuzuordnen sind (Kajava et al., 2025; siehe nachfolgender Abschnitt). Indem sich die digitale Elite zugleich als moralisch verantwortungsbewusst inszeniert und auf die Cité des Staatsbürgers rekurriert, etwa durch das Versprechen,

die transformative Kraft der Technologie dem Gemeinwohl zu widmen, gelingt es ihr, gesellschaftliche Zustimmung zu mobilisieren (Kajava et al., 2025; Nachtwey & Seidl, 2023).

Für die politische Praxis wird solutionistischen Narrativen insbesondere in Zeiten gesellschaftlicher Krisen und tiefgreifender Transformationen Plausibilität zugesprochen, indem sie Orientierung, Handlungsfähigkeit und technische Expertise miteinander verbinden. In Phasen erhöhter Unsicherheit, wie etwa während der Corona-Pandemie oder angesichts des Klimawandels, entsteht ein gesteigerter Bedarf nach schnellen und scheinbar rationalen Entscheidungen. Politische Institutionen greifen in solchen Situationen verstärkt auf Expert:innenwissen zurück, um Komplexität zu reduzieren und der Bevölkerung Sicherheit zu vermitteln (Bartl, 2023; Schopmans, 2022; Sharon, 2021; Zürn, 2018). Das solutionistische Narrativ bietet in diesem Kontext nicht nur eine technische Antwort auf gesellschaftliche Probleme, sondern auch eine Form politischer Legitimation: Es übersetzt Unsicherheit in steuerbare, technisch lösbare Aufgaben und verleiht politischem Handeln den Anschein wissenschaftlicher Objektivität. Die technologische Expertise fungiert dabei als epistemische Autorität, die politische Verantwortung teilweise substituiert und den Eindruck erweckt, Krisen ließen sich primär durch technische Innovation und datenbasierte Steuerung bewältigen (Obendiek & Seidl, 2023). So zeigte sich etwa während der Corona-Pandemie, dass politische Entscheidungsträger:innen auf digitale Instrumente wie Kontaktverfolgungs-Apps setzten, um schnell und effektiv zu reagieren (Bartl, 2023; Csernatoni, 2020; Martins et al., 2021; Milan, 2020; Sharma, 2022; Sharon, 2021). Morozov (2020) beschreibt diese Entwicklung als Übergang zu technologisch-solutionistischen Strategien als „Default Option“, um das politische Vakuum in Zeiten der Unsicherheit zu füllen. In solchen Momenten dient technologiegestützte Steuerung nicht nur der Krisenbewältigung, sondern auch als Legitimationsstrategie für politisches Handeln in einer Situation erhöhter Unsicherheit.

Während bislang nur vereinzelte empirische Studien dazu vorliegen, inwiefern politische Akteur:innen diese solutionistischen Narrative der Digital-Elite aufgreifen, verweisen die vorliegenden Erkenntnisse über die politische KI-Debatte in Deutschland darauf, dass die deutsche Regierung einer optimistisch-deterministischen Deutung einer durch KI geprägten Zukunft folgt: „Embedding AI in the institutional complex of the German welfare state comes with the utopian vision that AI, like social policy, operates as an emancipatory rather than a dominating or Orwellian force“ (Ossewaarde & Gülenç, 2020, S. 57). In der nationalen KI-Strategie der Bundesregierung (2018, 2020) wird KI etwa als Schlüsseltechnologie präsentiert, die zentrale gesellschaftliche Herausforderungen wie den demografischen Wandel, steigende Gesundheitskosten oder die Bewältigung der Coronapandemie adressieren soll (Hjaltalin & Sigurdarson, 2024; Ossewaarde & Gülenç, 2020). Köstler und Ossewaarde (2022) identifizieren in einer

systematischen Framinganalyse die Darstellung von „KI als Allheilmittel“ als einen der zentralen Deutungsrahmen des politischen Diskurses. KI erscheint darin als Technologie mit nahezu unbegrenztem Lösungspotenzial, während Risiken oder normative Konflikte in den Hintergrund treten. Diese Erzählung kann somit durchaus als solutionistische Deutung interpretiert werden, die technologische Innovationen als inhärent fortschrittlich darstellt und politische, soziale oder moralische Aushandlungen weitgehend ausblendet (Ossewaarde & Gülenç, 2020). Auch eine international vergleichende empirische Betrachtung zeigt, dass Regierungen den bestehenden KI-Hype eher verstärken, als ihn kritisch zu hinterfragen: „Recent technological advancements in AI are severely hyped, and governments contribute to this hype, instead of acting as critical watchdogs, soberly assessing the risks and potentials“ (Bareis & Katzenbach, 2022, S. 875).

6.1.1.2. Die Rechtfertigung Künstlicher Intelligenz im Kontext des internationalen Wettbewerbs und Wohlstandsversprechen

Darüber hinaus deuten empirische Befunde darauf hin, dass die politische Diskussion über den digitalen Wandel und insbesondere über KI an historisch etablierte Rechtfertigungsmuster anknüpft, die den Technologieeinsatz vor allem mit internationaler Wettbewerbsfähigkeit und der Sicherung wirtschaftlichen Wohlstands begründen. So beansprucht etwa die nationale KI-Strategie der Bundesregierung (2018, 2020) im Kontrast zur internationalen, insbesondere außereuropäischen Konkurrenz für Deutschland eine führende Rolle in der europäischen Technologiepolitik (Hirsch-Kreinsen & Krokowski, 2024b; Ossewaarde & Gülenç, 2020). Vor allem der Produktionssektor, traditionell das Rückgrat der deutschen Wirtschaft, steht dabei im Zentrum. Digitaltechnologien und KI gelten hier nicht nur als Instrument zur Steigerung der Wettbewerbsfähigkeit Deutschlands und Europas (Bareis & Katzenbach, 2022; Lambach & Oppermann, 2023), sondern als Schlüsseltechnologie zur Anpassung bestehender industrieller Strukturen an die Herausforderungen globaler Digitalisierung (Thelen, 2019). Die Notwendigkeit, auf den Vorsprung anderer Wirtschaftsmächte, allen voran der USA mit ihrem Innovationszentrum im Silicon Valley, aber ebenso Chinas oder Russlands, zu reagieren, prägt die politische Argumentation in Deutschland. Ausgelöst durch optimistisch bis utopische Erwartungen entsteht der Eindruck eines permanenten Aufholbedarfs und eines globalen Wettlaufs, in dem die erfolgreiche Implementierung von KI als Voraussetzung für ökonomische Stabilität und internationale Wettbewerbsfähigkeit gilt (Köstler & Ossewaarde, 2022; Ulnicane et al., 2022).

Das Framing von KI in der politischen Debatte erfolgt nicht nur als Allheilmittel für einzelne gesellschaftliche Herausforderungen, sondern zugleich als zentrale Zukunftstechnologie. Neben der Steigerung von Wettbewerbsfähigkeit wird insbesondere die Sicherung ökonomischen

Wohlstands betont, wobei dem KI-Einsatz Effizienz- und Effektivitätsgewinne zugeschrieben werden (Köstler & Ossewaarde, 2022). KI erscheint damit als unverzichtbare Voraussetzung für die gesellschaftliche und wirtschaftliche Zukunftsfähigkeit Deutschlands. Ihre Weiterentwicklung wird als notwendig dargestellt, um wirtschaftliche Stabilität, Sicherheit und den Erhalt bestehender Lebensstandards zu gewährleisten, während zugleich zukünftiger Wohlstand in Aussicht gestellt wird (Bareis & Katzenbach, 2022). Auf diese zugeschriebene Unausweichlichkeit verweisen auch Analysen internationaler bildungspolitischer Strategiepapiere. Sie zeigen, dass sowohl dem Erwerb von KI-Kompetenzen als auch dem Einsatz von KI im Bildungswesen eine zentrale Bedeutung für die Sicherung von Wettbewerbsfähigkeit und wirtschaftlichem Wachstum zugeschrieben wird (Nemorin et al., 2023; Schiff, 2022). Der durch diese Unausweichlichkeit skizzierte Handlungsdruck sieht die größte Gefahr des digitalen Wandels im Niedergang der deutschen Wirtschaft in Folge von Untätigkeit und mangelnder Antizipation internationaler Entwicklungen (Köstler & Ossewaarde, 2022). Dabei wird der Verlust technologischer Souveränität durch Abhängigkeiten von ausländischen Systemen als sicherheitspolitische Gefahr beschrieben (Lambach & Oppermann, 2023; Radu, 2021; Schopmans, 2022). Der Aufbau nationaler Schutz- und Vorsorgestrukturen wird entsprechend als Voraussetzung verstanden, um die Vision einer unausweichlich gewinnbringenden und souveränen KI-Nutzung zu verwirklichen (Köstler & Ossewaarde, 2022).

6.1.2. AI as „political-administrative necessity" – Mangelnde Politisierung aufgrund der Unausweichlichkeit Künstlicher Intelligenz

Die Sichtweise eines optimistischen Determinismus begründet eine weitreichende Autonomisierung der Technik, die gesellschaftliche Veränderungen hervorbringt, aufgrund ihrer dargestellten Unumgänglichkeit jedoch nur begrenzt politisch oder sozial gestaltbar erscheint. Technik wird als exogener Faktor betrachtet, der gesellschaftliche Entwicklung determiniert. Unabhängig davon, ob diese Deutung mit positiven Fortschrittserwartungen oder mit negativen Zukunftsszenarien verbunden ist, erzeugt sie eine deterministisch-fatalistische Ohnmacht, die politische Gestaltbarkeit systematisch einschränkt. Hay (2007) kritisiert diese Haltung, da sie Politisierung als notwendige Voraussetzung legitimer Entscheidungsfindung unterminiert:

> Fatalism and resignation are the antithesis of politics. The extent to which our destiny is determined by processes beyond our control is the extent to which it is non-political. (…). Yet, by the same token, when we question the inevitability of processes previously left to the benign stewardship of fate, and bring them once more under human influence and design, we expand the realm of politics, engaging thereby in a process of politicization. (Hay, 2007, S. 67)

In diesem Lichte äußerten sich bereits die verbreitete Technikbegeisterung der Nachkriegszeit sowie die Annahme der Eigenlogik von Technik in einer geringen Politisierung technologischer Entwicklungen. Konflikte über technologische Risiken oder Alternativen blieben in dieser Phase weitgehend ausgeblendet, da Technik als inhärent fortschrittlich und wertneutral galt. Die technische Gestaltung blieb, einem Top-down-Ansatz folgend, Expert:innen vorbehalten, während die Politik sich auf die Bereitstellung von Ressourcen für Forschung und Entwicklung, nicht aber auf die demokratische Aushandlung von Zielen oder Nebenwirkungen konzentrierte (Bogner, 2021b; Edler & Wittmann, 2024; Saretzki, 2001; Schubert, 1991; Weyer, 2008). Damit verfestigte sich ein technokratisches Selbstverständnis, das politische Steuerung auf die Förderung wissenschaftlich-technischen Fortschritts reduzierte, welcher als inhärente Quelle gesellschaftlichen Fortschritts galt.

Diese historische Konstellation liefert eine aufschlussreiche Analogie für die gegenwärtige KI-Debatte. Indem technologische Entwicklungen als naturwüchsige oder gar unvermeidbare Prozesse dargestellt werden, kann das positiv gezeichnete Bild von KI als Motor gesellschaftlichen Fortschritts als Versuch einer Depolitisierung interpretiert werden. Dabei werden Risiken und Herausforderungen vorwiegend als technische Fragen dargestellt, deren Lösung primär in die Hände der Wirtschaft gelegt wird. Dies verstärkt die Wahrnehmung von KI als „political-administrative necessity“ (Köstler & Ossewaarde, 2022, S. 257), deren Umsetzung die Disziplin und Zustimmung aller gesellschaftlichen Akteur:innen erfordert. Kritische Narrative des digitalen Wandels scheinen hingegen nur vereinzelt im politischen Diskurs zum Ausdruck zu kommen und äußern sich eher unspezifisch, beispielsweise als diffuse Bedenken gegenüber einer potenziellen Aushöhlung demokratischer Strukturen durch KI (Lambach & Oppermann, 2023). Des Weiteren deutet auch die im Vergleich zu anderen Politikfeldern eher untergeordnete Rolle der Digitalpolitik auf ein geringes Ausmaß der Politisierung hin, wenngleich ein quantitativer Anstieg digitalpolitischer Schwerpunkte im deutschen politischen Diskurs zu verzeichnen ist (Dumbach et al., 2023; König & Wenzelburger, 2019). Während Bao et al. (2022) auf erste Anzeichen einer beginnenden Politisierung von KI in der US-amerikanischen Öffentlichkeit hinweisen, gehen Siewert und König (2019) in ihrer Analyse der Salienz der Digitalisierung in deutschen Parteiprogrammen davon aus, dass die Depolitisierung der Digitalisierung fortbestehen werde, solange keine externen Einflüsse auf das Politikfeld einwirken.

6.2. Der Kontrollverlust der Technik: Die Perspektive des pessimistischen Determinismus

Externe Einflüsse sind dabei auch in der historischen Betrachtung als Auslöser einer zunehmenden Technologiepolitisierung sowie eines Wandels der Technikdeutung hin zu einem

pessimistischen Determinismus zu identifizieren. So wuchs ab Mitte der 1960er Jahre vor dem Hintergrund zunehmender negativer gesellschaftlicher Technikerfahrungen das öffentliche Bewusstsein für Technikrisiken und unintendierte Folgen. Auslöser hierfür waren unter anderem der Einsatz von Militärtechnologien sowie folgenreiche technische Fehlschläge wie die Reaktorkatastrophen von *Three Mile Island* (1979) und Tschernobyl (1986) oder die Explosion der Raumfähre *Challenger* (1986). Die Legitimität des Technikeinsatzes geriet durch diese Ereignisse ebenso ins Wanken wie das Vertrauen in die staatliche Steuerungsfähigkeit, da die technischen Fehlschläge als Versagen der staatlich-technokratischen Forschungsförderung interpretiert wurden (Bogner, 2021b; Brown, 2003; Degele, 2002; Geels & Verhees, 2011; Weyer, 2008). Technik wurde nicht länger als neutraler Fortschrittsmotor wahrgenommen, sondern als potenzielle Quelle von Bedrohung und Kontrollverlust. Damit rückten in einer zunehmend kritischen techniksoziologischen Betrachtung Fragen nach der Zweckmäßigkeit technischer Entwicklungen in den Vordergrund (Horkheimer & Adorno, 1947/1969).

Im Zentrum des sich herausbildenden pessimistischen Determinismus stand die Vorstellung einer negativen Eigendynamik der Technik. Während im optimistischen Paradigma technische Rationalisierung mit Fortschritt und gesellschaftlichem Nutzen gleichgesetzt wurde, betonten Autoren wie Ellul (1954/2021), Heidegger (1962/2011) oder Winner (1977/2001) die destruktiven Konsequenzen einer sich selbst verstärkenden autonomen Technik, die sich der gesellschaftlichen Kontrolle entzieht. Zusätzliche Bedeutung gewinnen diese Auffassungen im Zuge einer neuen Qualität von Technik, die mit der Entwicklung moderner Informations- und Algorithmisierungstechnologien (z. B. des Turing-Rechners, der als heutiger Grundstein von KI-Anwendungen gilt) einhergeht. Technik begann, auch kognitive menschliche Leistungen zu ersetzen, und verlor damit ihren Charakter als lediglich den Menschen entlastende Maschine (Gehlen, 1986/2017; Passoth, 2008). Sie wurde verstärkt als Instrument der Überwachung und gesellschaftlichen Kontrolle kritisiert (Foucault, 1975/2024). Diese Entgrenzung des Technischen verstärkte in der techniksoziologischen Kritik die Wahrnehmung einer unkontrollierbaren Dynamik, in der erstens technische Innovationen fortlaufend neue technische Lösungen hervorbringen und damit eine sich selbst legitimierende Rationalisierungskette erzeugen, zweitens gesellschaftliche Zwänge entstehen, die Handlungsspielräume einschränken, und drittens ein menschlicher Kontrollverlust gegenüber der Technik eintritt (Ellul, 1954/2021; Freyer, 1987; Gehlen, 1957/2004; Horkheimer, 1947/1967; Horkheimer & Adorno, 1947/1969; Marcuse, 1941; Schelsky, 1961; Winner, 1977/2001).

Wie zuvor bereits erwähnt, spannen sich deterministische Technikdeutungen dabei entlang eines Kontinuums auf (Dafoe, 2015; Marx & Smith, 1994). Ropohl (2016) differenziert hier

zwischen einem genetischen und einem konsequenziellen Determinismus. Ersterer kommt beispielsweise in der Beschreibung Elluls (1954/2021) einer Unaufhaltsamkeit von Technik zum Ausdruck und ist in diesem Sinne der Technik inhärent. Der konsequenzielle Determinismus betont dabei stärker die Prägung sozialen Handelns durch Technik und beschreibt die gesellschaftlich folgenreichen Technikwirkungen nicht selten als Konsequenz der Planungseuphorie der Moderne (z. B. Schelsky, 1961; Ogburn, 1969).

Während ein genetischer Determinismus zunehmend als überwunden gilt (M. Bauer, 1995; Bieber & Deiß, 1997; Katzenbach, 2018; Weyer, 2008), erscheinen konsequenzielle Warnungen vor einer Eigendynamik technischer Entwicklungen weiterhin aktuell (R. Boyd & Holton, 2018; Dafoe, 2015; Mai, 2011). Dies zeigt sich sowohl im Wiederaufleben technokratiekritischer Positionen (siehe Abschnitt 6.2.1) als auch in der anhaltenden Debatte über den transformativen Charakter neuer Technologien mit ihren ambivalenten Folgen zwischen Fortschritt und Risiko. So weisen einige Autor:innen auf den disruptiven Charakter von KI-Anwendungen hin, wenn sie beispielsweise die Transformation der Arbeitswelt und den damit prognostizierten Verlust von Arbeitsplätzen oder den Wandel des Bildungswesens in den Blick nehmen (Ford, 2015; Moravec & Martínez-Bravo, 2023; Păvăloaia & Necula, 2023).

In der gegenwärtigen KI-Debatte finden sich neben deterministischen Fortschritts- und Wohlstandsversprechen auch kritische Reflexionen, teils sogar von prominenten Vertreter:innen der Tech-Industrie selbst: So gehört Elon Musk zu den Erstunterzeichner:innen des offenen Briefes „Pause Giant AI Experiments“, in dem im Frühjahr 2023 ein sechsmonatiges Moratorium für die Weiterentwicklung leistungsstarker KI-Systeme gefordert wurde. Dieser Appell versammelte bereits nach wenigen Tagen über 1.000 Unterstützer:innen (Future of Life Institute, 2023). In dem offenen Brief wird vor tiefgreifenden gesellschaftlichen Risiken gewarnt, die mit der KI-Entwicklung „menschlich-kompetitiver“ Intelligenz einhergehen:

> Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: (…) *Should* we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? *Should* we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. (Future of Life Institute, 2023; Herv. i. O.)

Die in diesem Appell formulierte Forderung nach einem sechsmonatigen Moratorium solle der Politik die nötige Zeit verschaffen, um auf diese Risiken zu reagieren und geeignete regulatorische Maßnahmen zu ergreifen. Zudem betonen sie die Notwendigkeit erheblicher institutioneller und finanzieller Ressourcen, um auf die erwarteten, durch KI ausgelösten Disruptionen

der Ökonomie und insbesondere der Demokratie angemessen reagieren zu können (Future of Life Institute, 2023). Die kausale Verknüpfung der KI-Entwicklung mit einem drohenden Kontrollverlust über die menschliche Zivilisation kann dabei als Ausdruck eines (konsequenziellen) Determinismus verstanden werden, in dem Technologie als Haupttreiber gesellschaftlicher Transformation gedacht wird. KI wird hier sowohl als Bedrohung als auch als zu regulierendes Risiko verstanden. Damit knüpft die hier vertretene KI-Deutung aus Sicht des pessimistischen Determinismus an zentrale Muster vergangener Technikkritik an.

6.2.1. Technikkritik: Das Legitimationsdefizit durch technischen Kontrollverlust und Technokratisierung

Historisch betrachtet führte ab Mitte der 1960er Jahre die wachsende Ernüchterung über negative Technikerfahrungen und die deterministische Annahme technischer Zwänge zu einer umfassenden Technik- und Technokratiekritik, die sich auf zwei zentralen Ebenen artikulierte und auf die Cité des Staatsbürgers verweist. Zum einen wurde der durch die vermeintliche Eigendynamik der Technik ausgelöste Kontrollverlust kritisiert, der als ein Widerspruch des Ideals mündiger Bürger:innen interpretiert werden kann. Die zunehmende Technisierung gesellschaftlicher Bereiche führe, so Habermas (1968), zu einem „quasi-autonome[n] Fortschritt von Wissenschaft und Technik“ (S. 77), der zunächst in einer Entpolitisierung der Bevölkerung und schließlich in einer Technokratisierung des Staates münde. Zum anderen verband sich diese Kritik mit einer antikapitalistischen Perspektive, die Technik als Ausdruck und Instrument ökonomischer Verwertungslogiken verstand. Die technokratische Reduktion politischer Entscheidungen auf Effizienz und Funktionalität wurde dabei als Widerspruch zu demokratischen Prinzipien öffentlicher Legitimation und Teilhabe interpretiert (Häußling, 2019; Weyer, 2008).

Die Vorstellung einer eigendynamischen Technik, die gesellschaftliche und politische Ziele zunehmend vorgibt, schränkt den politischen Entscheidungsspielraum ein und mindert den wahrgenommenen Legitimationsbedarf. Im Zuge dessen kritisiert Habermas (1968), ebenso wie Schelsky (1961), dass das Wohlstandsversprechen der Technik keiner Legitimation bedürfe, solange dieses unwidersprochen bliebe: „Die moderne Technik bedarf keiner Legitimität; mit ihr ‚herrscht‘ man, weil sie funktioniert und solange sie optimal funktioniert“ (Schelsky, 1961, S. 25). Diese Diagnose stellt somit die Rechtfertigungen des optimistischen Determinismus infrage und betont zugleich die Notwendigkeit kritischer Perspektiven.

Sowohl der befürchtete Kontrollverlust als auch die Kritik einer fortschreitenden Technokratisierung bilden auch gegenwärtig zentrale Elemente der KI-Kritik. In techniksoziologischen Debatten äußern sie sich etwa in der Sorge vor datenbasierter Überwachung und der Aushöhlung

demokratischer Entscheidungsprozesse. Damit knüpft die Kritik an die historischen Konfliktlinien der Technikdebatte an und spitzt diese im Zuge der digitalen Entwicklung neu zu.

6.2.1.1. Künstliche Intelligenz als Instrument der Kontrolle und Überwachung

In der techniksoziologischen Debatte betrifft ein zentraler Kritikpunkt der Verbreitung von KI ihre potenziell demokratiegefährdenden Effekte durch datenbasierte Kontrolle und Überwachung. Digitale Daten bilden die zentrale Ressource, die den Einsatz von KI erst ermöglicht. Ihre massenhafte Sammlung und Nutzung, insbesondere durch große Technologiekonzerne, aber beispielsweise auch durch Bildungsinstitutionen, führt zu neuen Formen digitaler Abhängigkeit, eingeschränkten Persönlichkeitsrechten und Machtasymmetrie (Cohen, 2019; Feldstein, 2019; Khanal et al., 2025; Prinsloo & Slade, 2015; Slade & Prinsloo, 2013). Persönliche Daten werden scheinbar freiwillig im Austausch gegen Effizienz, Bequemlichkeit und Zugang zu digitalen Diensten preisgegeben. Bei genauerer Betrachtung erweist sich diese „Freiwilligkeit“ jedoch als trügerisch (boyd & Crawford, 2012): Nutzer:innen fehlt es häufig an realen Alternativen sowie an Transparenz über die Datennutzung. Die Nutzung zentraler digitaler Dienste setzt meist die Zustimmung zur Datenerhebung voraus, sodass eine echte Entscheidungsfreiheit kaum besteht. Die Datenfreigabe wird somit zur zwingenden Voraussetzung zur digitalen Teilhabe (Yeung, 2018).

Die algorithmische Verarbeitung, beispielsweise im Rahmen einer KI-basierten Kreditwürdigkeitsprüfung, erfolgt vielfach ohne aktive Zustimmung der Betroffenen. Dadurch verlieren Individuen die Kontrolle darüber, welche Daten zu welchen Zwecken verarbeitet werden (Davidow, 2014). Dies kann als Kritik der Cité des Staatsbürgers gedeutet werden, indem zentrale demokratische Werte wie Selbstbestimmung unterminiert werden. In der Folge wird zunehmend vor einer KI-basierten gesellschaftlichen Überwachung gewarnt, die die Grundlagen demokratischer Freiheit gefährdet (Schopmans, 2022; Zuboff, 2019). So sieht Hesse (2020) nicht nur die Gefahr einer Entmündigung des Einzelnen, sondern ebenso „einer sukzessiven Aushöhlung von Staaten und demokratischen Institutionen“ (S. 341). Demokratien laufen damit Gefahr, ihre Steuerungsfähigkeit an anonyme Akteur:innen und Technologieunternehmen zu verlieren. Denn in Anbetracht einer „unaufhaltsam voraneilenden Technik und Wissenschaft einerseits und einer willig folgenden Politik und Wirtschaft andererseits stellt die Sicherung der Demokratie und ihre Bewahrung vor gefährlichen Auflösungstendenzen eine gewaltige Herausforderung dar – vielleicht sogar die zentrale des laufenden Jahrhunderts“ (Hesse, 2020, S. 345). Eine Gegenmaßnahme erfordert daher die Ausbildung humanistischer Bildungsideale, insbesondere des kritischen Denkens des Einzelnen, sowie eine umfassende Politisierung der Technologiedebatte im Sinne einer breiten gesellschaftlichen Diskussion (Hesse, 2020).

6.2.1.2. Das technokratische Repräsentationsdefizit künstlich intelligenter Entscheidungen

Darüber hinaus richtet sich die techniksoziologische KI-Kritik zunehmend auf die durch datenbasierte Steuerung erzeugte asymmetrische Informationsordnung, in der Individuen beobachtet und bewertet werden, ohne selbst Einblick in die Funktionsweise und Entscheidungsgrundlage der Systeme zu erhalten. In der Folge verschiebe sich das Machtverhältnis zulasten demokratischer Prinzipien wie Rechenschaftspflicht, Öffentlichkeit und rechtlicher Widerspruchsmöglichkeit. Bürger:innen verlieren zunehmend die Möglichkeit, algorithmisch getroffene Entscheidungen, die ihr Leben direkt betreffen, zu verstehen oder anzufechten (Yeung, 2018). Damit unterlaufen diese Systeme zentrale Anforderungen demokratischer Rechenschaft, beruhen liberal-demokratische Ordnungen doch auf der Idee, dass Bürger:innen unter Regeln leben, an deren Formulierung sie mitwirken können und deren Anwendung überprüf- und anfechtbar bleibt (Waldron, 1987). Diese doppelte Transparenz der demokratischen Teilhabe an der Regelsetzung und der rechtlichen Kontrolle der Regelanwendung bildet eine wesentliche Grundlage politischer Autorität. Aufgrund der Intransparenz des Entwicklungsprozesses KI-basierter Systeme einerseits sowie der von ihnen getroffenen Entscheidungen andererseits laufen entsprechende Entwicklungen jedoch Gefahr, dieses Gleichgewicht zu stören (Hildebrandt, 2016; Pasquale, 2015).

Die Intransparenz und Exklusion aus dem Entscheidungsprozess münden somit in einem technokratischen Repräsentationsdefizit (Li, 2024; Ma, 2020; Salais, 2016), wie es bereits in der Technikdebatte der 1970er und 1980er Jahre problematisiert wurde.[14] Während sich die damalige Kritik vor allem gegen die „Anpassung der Politik an die Rationalität der Technik“ (Mai, 2011, S. 145) richtete, zielt die heutige Kritik zunehmend auf den technokratischen Einfluss der Technologie-Konzerne: „Die technischen Strukturen und Systeme, die Kultur und Gesellschaft prägen, (…) sind weitgehend die Folge privatwirtschaftlicher Initiativen, der Aushandlungsergebnisse von Expertennetzwerken und individueller Entscheidungen. Geblieben ist von der klassischen Technokratie die Entpolitisierung der Entscheidungen“ (Mai, 2011, S. 145). Der Einfluss der *Big-Tech*-Konzerne beschränkt sich dabei nicht ausschließlich auf die technische Entwicklung, sondern zunehmend auch auf die politische Gestaltung und Steuerung von KI (Benkler, 2019; Khanal et al., 2025; Lindman et al., 2023). Zugleich trägt die Hinwendung zu

[14] Die Technokratiekritik wird dabei nicht ausschließlich aus einer deterministischen Perspektive formuliert. Während sie einerseits an konsequenziell-deterministische Annahmen anschließt, die auf eine Eigendynamik technischer Entwicklungen und deren politische Entkopplung verweisen, wird sie andererseits ebenso aus sozialkonstruktivistischen Perspektiven geäußert, die die Exklusion gesellschaftlicher Akteur:innen und die Verengung politischer Aushandlungsprozesse kritisieren. Passoth (2008) beschreibt die Technokratie-Debatte vor diesem Hintergrund als eine der wenigen direkten Auseinandersetzungen zwischen deterministischen und sozialkonstruktivistischen Standpunkten in der Techniksoziologie.

einem an gesellschaftlicher Nachfrage orientierten *Demand-Pull*-Modell der Technologiepolitik zu einer veränderten Vorstellung des Verhältnisses von Gesellschaft, Politik und Technik bei. So verstärken nicht nur die Macht der Technologieunternehmen, sondern auch eine unkritische gesellschaftliche Nachfrage nach digitalen Anwendungen die technokratischen Tendenzen, anstatt eine Politisierung der KI-Entwicklung zu fördern (Mai, 2011). Während diese Risiken in der techniksoziologischen Debatte zunehmend diskutiert werden, bleibt bislang offen, inwiefern sie Widerhall in der politischen Debatte finden.

6.2.2. Der rekursive Politisierungsprozess: Technikkritik als Ausdruck reaktiver Politisierung und als Folge technischer Fehlschläge

Die historische Betrachtung der Technikkritik erlaubt zwei Schlussfolgerungen: Zum einen kann Technikkritik als Ausdruck einer reaktiven Politisierung verstanden werden, deren Mangel in derselben Kritik noch beanstandet wird. Zum anderen führt sie zu einer zunehmenden Polarisierung der öffentlichen Technikwahrnehmung. Ab Mitte der 1960er Jahre führte die Kritik an Technik und Technokratie nach ersten technischen Fehlschlägen zu einer verstärkten öffentlichen Auseinandersetzung über die Folgen technischer Entwicklungen (Geels & Verhees, 2011; Weyer, 2008). Diese reaktive Politisierung zeigte sich etwa in den globalen Protesten der 68er-Bewegung gegen technische Kontrollzwänge und die Steuerung technokratischer Eliten (Feenberg, 1999) sowie in politischen Initiativen zur wissenschaftlichen Bewertung technologischer Risiken, wie der ersten *Deutschen Risikostudie Kernkraftwerke* (Gesellschaft für Reaktorsicherheit, 1980; Passoth, 2008; Saretzki, 2001).

Mit der wachsenden Kritik wich die anfängliche Technikgläubigkeit einer zunehmenden Polarisierung der öffentlichen Technikwahrnehmung (Schubert, 1991). Somit zeigt sich die Ambivalenz zwischen utopischen und dystopischen technikdeterministischen Zukunftsvisionen: „To optimists, such a future is the outcome of many free choices and the realization of the dream of progress; to pessimists it is a product of necessity's iron hand, and it points to a totalitarian nightmare" (Marx & Smith, 1994, S. xii). Eine Polarisierung, die zunehmend auch der Auseinandersetzung mit KI zugesprochen wird, wobei sich technikoptimistische Narrative, die KI als Allheilmittel stilisieren, und technikpessimistische Warnungen vor Kontrollverlust, Überwachung und Entdemokratisierung gegenüberstehen (Callaghan, 2018; Diaz-Bone & Schrör, 2022; Gallego & Kurer, 2022; Katzenbach & Ulbricht, 2019; Morozov, 2018).

Gerade pessimistische Positionen formieren sich häufig als Reaktion auf enttäuschte Erwartungen, die nach dem Sichtbarwerden negativer Auswirkungen von KI entstehen. Technische Fehlschläge und unerwartete Nebenfolgen lösen öffentliche Kritik aus und führen damit *ex post* zu

einer Politisierung von KI. Ulbricht (2020) zeigt in mehreren Fallstudien[15], dass die Einführung von KI-Systemen zunächst im Sinne eines optimistischen Determinismus als neutrale, effiziente Werkzeuge gerechtfertigt wurden und dadurch depolitisiert erschienen. Erst wenn negative Folgen des KI-Einsatzes sichtbar werden, werden sie zum Gegenstand öffentlicher Kontroversen. Zu diesen Fallbeispielen zählt der AMS-Algorithmus des österreichischen Arbeitsmarktservice, der Arbeitssuchende nach ihren prognostizierten Integrationschancen klassifizierte. Das System versprach Effizienz und Objektivität, stieß jedoch auf erhebliche Kritik, da bereits die Klassifizierung der Arbeitsvermittelbarkeit problematisiert wurde und sich zudem zeigte, dass Frauen auf Grundlage historischer Daten systematisch schlechter bewertet wurden. Dadurch reproduzierte die Software bestehende Ungleichheiten, anstatt sie abzubauen (Allhutter et al., 2020; Ulbricht, 2020). Dieses Beispiel verdeutlicht den von Ulbricht (2020) skizzierten dreistufigen rekursiven Politisierungsprozess von KI:

(1) Depolitisierung durch KI: Zu Beginn führt die algorithmische Formalisierung sozialer Entscheidungen zu einer Entpolitisierung normativer Fragen. Indem Entscheidungen in technische Modelle und Software übersetzt werden, werden Kontingenzräume geschlossen und politische Werturteile als technische Parameter verborgen.
(2) Politisierung von KI: Mit dem Sichtbarwerden diskriminierender Effekte oder Fehlurteile einer KI-Anwendung entzündet sich Kritik an der vermeintlichen Objektivität der Technologie, deren Legitimität in der Folge infrage gestellt wird.
(3) Repolitisierung sozialer Normen: Schließlich weitet sich die Auseinandersetzung auf die zugrunde liegenden Wertentscheidungen der KI aus. Es entsteht eine Debatte darüber, welche sozialen Normen und Gerechtigkeitsvorstellungen in technische Systeme eingeschrieben sind.

Die reaktive Politisierung der KI erschüttert damit das Fortschrittsversprechen des optimistischen Determinismus, das KI als neutrales und objektives Problemlösungsinstrument deutet, und schwächt zugleich die epistemische Autorität derjenigen, die dieses Versprechen vertreten. Zugleich bringt sie Fragen der Gemeinwohlorientierung und Gerechtigkeit zurück in den politischen Raum und erhöht gleichzeitig den Rechtfertigungsdruck der politischen Entscheider:innen. Vormals verborgene normative Annahmen werden sichtbar, was neue Aushandlungsräume und konkurrierende Rechtfertigungsnarrative unterschiedlicher Policy-Ziele eröffnet (Taeihagh, 2021). Dieser rekursive Politisierungsprozess – von der anfänglichen

[15] Untersucht wurden das US-amerikanische COMPAS-System aus dem Bereich der Strafverfolgung, der österreichische AMS-Algorithmus im Kontext des Arbeitsmarkts sowie Zalandos Zonar zur Bewertung von Mitarbeitenden und der Algorithmus der Schufa, der zum Zwecke der Bonitätsprüfung eingesetzt wird.

Depolitisierung über die reaktive Politisierung der Technologie bis hin zur Repolitisierung ihrer moralischen Grundlagen – verdeutlicht, dass deterministische Deutungen von KI nicht haltbar scheinen. Die Infragestellung der Einhaltung sozialer Normen verweist auf die soziale Konstruiertheit von KI-Systemen und markiert damit den Übergang zu einem sozialkonstruktivistischen Verständnis von Technik.

6.3. Die Demokratisierung technischer Gestaltung als Folge der Anerkennung der Ambivalenz von Technik: Die Perspektive des Sozialkonstruktivismus

Ab den 1980er Jahren verschiebt sich das techniksoziologische Verständnis zunehmend vom technikdeterministischen zu einem sozialkonstruktivistischen Paradigma. Angesichts der zunehmenden Technikkritik und Kontroversen um Atomkraft und Gentechnik wuchs das Bewusstsein für die Ambivalenz technischer Entwicklungen. Es eröffnete sich ein Spannungsfeld zwischen Effizienzversprechen und Kontrollverlust, gesellschaftlichem Nutzen und Risiko (Bogner, 2021b). Der Fokus verlagerte sich von der Frage, *was* Technik ist, hin zur Frage, *wer* unter *welchen Bedingungen* ihre Entwicklung und Nutzung beeinflusst (Rammert, 1992). Technik wurde nicht länger als autonome Kraft des Fortschritts oder als Auslöser von Zwängen, sondern als Ergebnis sozialer und politischer Aushandlungsprozesse verstanden (Pinch & Bijker, 1984; Weyer, 2008).

Empirische Beobachtungen begründeten diesen Paradigmenwechsel: Einflussreiche Studien belegen, dass technologische Entwicklungen vom Fahrrad bis hin zur Atomkraft Ergebnisse komplexer sozialer Aushandlungen sind. Unterschiedliche Akteur:innen mit divergierenden Interessen beeinflussen dabei nicht nur die Entstehung, sondern auch die Nutzung und Bedeutung technischer Artefakte, die je nach Kontext zu verschiedenen Zwecken eingesetzt werden können (Bijker, 1995; Linde, 1972; Noble, 1978; Passoth, 2008; Pinch & Bijker, 1984; Weyer, 2008). Diese Erkenntnis begründet ein neues Kontingenzbewusstsein, das der Idee einer unkontrollierbaren Eigendynamik der Technik eine Absage erteilt, sodass ein strenger Technikdeterminismus in der wissenschaftlichen Debatte fortan als überwunden gilt (Bauer, 1995; Bieber & Deiß, 1997; Katzenbach, 2018; Weyer, 2008), wenngleich bereits aufgezeigt wurde, wie dieser in gegenwärtigen KI-Debatten weiterhin zum Ausdruck kommt. Das sozialkonstruktivistische Paradigma rückt damit die soziale Gestaltbarkeit und Interpretationsbedürftigkeit technischer Entwicklungen ins Zentrum (Brown, 2015; Doezema & Frahm, 2023; Pinch & Bijker, 1984; Rammert, 1992).

6.3.1. Die Anerkennung der politischen Qualität und Ambivalenz von Technik

Diese Erkenntnis erwuchs neben der sichtbaren Ambivalenz, unter anderem hervorgehoben durch die Konfrontation technikoptimistischer und technikpessimistischer Technikdeutungen,

auch aus dem Bewusstsein für die politische Qualität von Technik. In seinem einflussreichen Beitrag „*Do Artifacts Have Politics?"* legt Winner (1980) dar, dass technische Artefakte soziale Machtverhältnisse reproduzieren oder neu begründen können. Am Beispiel absichtlich niedrig gebauter Brücken New Yorks zeigt Winner, wie technische Gestaltung soziale Exklusion herbeiführen kann, indem sie bestimmten Bevölkerungsgruppen den Zugang zum öffentlichen Raum verwehrt. So bewirkte die Niedrigbaukonstruktion, dass Busse, die damals als primäres Fortbewegungsmittel sozial benachteiligter und insbesondere dunkelhäutiger Bevölkerungsgruppen dienten, diese nicht passieren konnten. PKWs als bevorzugtes Transportmittel wohlhabenderer Schichten konnten hingegen problemlos unter den Brücken hindurchfahren. Auch gegenwärtige Beispiele wie die algorithmische Diskriminierung durch das Prognosesystem COMPAS, das Rückfallrisiken für Straftäter:innen systematisch zuungunsten dunkelhäutiger Personen bewertet (Angwin et al., 2016), verdeutlichen die politische Reichweite technischer Artefakte. Technik erweist sich damit nicht als neutral, sondern als strukturierende Kraft mit politischer Reichweite, die jedoch nicht naturwüchsig erscheint, sondern das Ergebnis sozialer Konstruktionsprozesse ist (Häußling, 2019; Katzenbach, 2018). Sie kann, wie Winner (1980) darlegt, ähnlich wie Gesetze langfristige gesellschaftliche Ordnungen etablieren.

Diese Einsicht in die politische Qualität technischer Artefakte gewinnt im Kontext von KI besondere Bedeutung, da KI unmittelbar in soziale und politische Prozesse eingreift. Ihre Logik der Quantifizierung zielt auf die Messbarmachung von Verhalten und Leistung und wird so zu einem Instrument der Regulierung, Kontrolle und Legitimation politischer Maßnahmen. Im Rahmen neoliberaler Steuerungslogiken wird Quantifizierung damit zu einem zentralen Mittel administrativer und staatlicher Macht (Desrosières, 1998, 2014/2025; Fourcade & Healy, 2024; Li, 2024). KI erscheint dabei nicht nur als technologische Innovation, sondern als Fortsetzung und Zuspitzung eines gesellschaftlichen Optimierungsimperativs. Die algorithmische Verarbeitung individueller Daten im Dienste effizienter Entscheidungsfindung steht in direkter Kontinuität zur Metrisierung und Standardisierung sozialer Prozesse in neoliberalen Gesellschaften (Desrosières, 2014/2025; Diaz-Bone & Didier, 2016; Espeland & Stevens, 2008; Lünich, 2022). Die mit der Quantifizierung verknüpfte Messbarmachung von Leistung und die darauf aufbauende Bewertung von Effizienz werden zunehmend zur legitimatorischen Grundlage politischen Handelns:

> In addition to being true, neutral and indisputable, figures have become indispensable to the smooth running of institutions: without them, it is impossible to measure, control and justify the decisions taken. In the political sphere, numbers have thus acquired the role of instruments for coordinating actions. (Henneguelle, 2023, S. 12)

Gerade durch diese politische Dimension gewinnt die Ambivalenz technologischer Innovationen besondere Bedeutung. Verfahren der Quantifizierung können einerseits zur Sichtbarmachung sozialer Ungleichheiten und zur Artikulation von Gerechtigkeitsforderungen beitragen, indem sie gemeinsame Bewertungsmaßstäbe schaffen (Porter, 1995). Im Bildungsbereich etwa ermöglichen KI-gestützte Prognosen von Studienabbrüchen, strukturelle Benachteiligungen zu identifizieren und beispielsweise gezielte Unterstützungsmaßnahmen für gefährdete Studierende bereitzustellen. In diesem Sinne entfaltet die KI-Anwendung ein emanzipatorisches Potenzial (Ma, 2020). Andererseits bergen dieselben Systeme das Risiko, bestehende Ungleichheiten zu verfestigen, wenn sie zur Kontrolle individueller Leistungen beitragen und als Instrumente der Disziplinierung oder Externalisierung politischer Verantwortung eingesetzt werden (Danaher, 2016; Feenberg, 2002; Fourcade & Healy, 2024; Kelty, 2014; Li, 2024; Schiff, 2021). Durch die KI-basierte Metrifizierung von Leistung oder Verhalten erfolgt eine statistische Kategorisierung von Personengruppen, die soziale Hierarchien zwischen „besseren“ und „schlechteren“ Studierenden erzeugt. Diese Hierarchisierung kann wiederum (hochschul-)politische Entscheidungen legitimieren, etwa darüber, wer Zugang zu bestimmten Förderprogrammen erhält und wer damit welche Chancen erhält. KI kann somit nicht nur emanzipatorisch wirken, sondern ebenso bestehende Ungleichheiten reproduzieren und verstärken (Desrosières, 2015; Henneguelle, 2023; Ma, 2020; Mau, 2020).

6.3.2. Die techniksoziologische Kritik der Verschleierung der sozialen Konstruktion von Technik

In dem Bewusstsein der sozialen Gestaltbarkeit und politischen Qualität moderner Technologien richtet sich die Kritik im sozialkonstruktivistischen Verständnis zunehmend gegen den mangelnden Einbezug gesellschaftlicher Perspektiven in technologische Entwicklungsprozesse. Historisch wurde diese Kritik bereits im Zuge der Technikdebatten der 1970er und 1980er Jahre sichtbar. So artikulierten soziale Bewegungen gegen Atomkraft oder Gentechnik auch vor dem Hintergrund der Technokratiekritik (siehe Abschnitt 6.2.1) die Gefahr der Unterwanderung demokratischer Grundrechte und Entscheidungsprozesse sowie die Sorge um gesundheitliche und ökologische Risiken (Bogner, 2021b; Heyen et al., 2024; Rahm, 2023a), wie sie als Kritik der Cité des Staatsbürgers und der Ökologie verstanden werden können. Gleichzeitig wurde die wissenschaftliche Autorität infolge des Dissenses und des Sichtbarwerdens der Unsicherheit wissenschaftlicher Erkenntnis infrage gestellt (Bogner, 2021b; Wehling, 2021; Weyer, 2008). Im Zuge dessen verdeutlicht etwa Becks (1986/2016) Diagnose der Risikogesellschaft, dass der Umgang mit den durch die Gesellschaft selbstproduzierten und zwischen sozialen Gruppen ungleich verteilten technologischen Risiken zu einer zentralen

gesellschaftlichen Herausforderung wird (vgl. Luhmann, 1991). Dabei könne der Umgang mit den Risiken nicht exklusiv durch wissenschaftliche Expertise bewältigt werden, sondern erfordere vielmehr partizipative Aushandlungsprozesse (Beck, 1986/2016).

Vor diesem Hintergrund erscheint die heutige techniksoziologische KI-Kritik als Fortsetzung der staatsbürgerlichen Technikkritik vergangener Zeiten unter veränderten Bedingungen. Ein zentraler Fokus liegt auf der Verborgenheit der sozialen Konstruktion von Technik, die sowohl auf deterministische Technikdeutungen als auch auf Intransparenz und Exklusion gesellschaftlicher Akteur:innen aus zentralen Entscheidungsprozessen zurückzuführen ist. Kritisiert wird einerseits das Fortschritts- und Wohlstandsversprechen von KI als Ausdruck eines technikdeterministischen Denkens, das die politische Gestaltbarkeit der Technologie verkennt und wissenschaftliche Erkenntnisse über ihre soziale Konstruiertheit ignoriert (R. Boyd & Holton, 2018; Hay, 2007; Spencer, 2017). Andererseits bleiben die Aushandlungsprozesse und die ihnen zugrunde liegenden Wertentscheidungen, etwa darüber, welche Ergebnisse einer KI als wünschenswert gelten oder welche Risiken akzeptabel sind, meist im technischen Design verborgen und entziehen sich der öffentlichen und politischen Reflexion. So werden die Komplexität und die Kontingenz der damit verbundenen Entscheidungen für große Teile der Gesellschaft unsichtbar (Li, 2024). Die anfängliche Aushandlung von Konventionen gerät in Vergessenheit, während der Output der Systeme als objektiv und naturgegeben erscheint. Dadurch werden normative Grundlagen verschleiert, Machtungleichgewichte stabilisiert und demokratische Prinzipien untergraben (Desrosières, 1998; Henneguelle, 2023; Lazar, 2024; Ma, 2020; O'Neil, 2016; Salais, 2016; Weinberg, 2022).

Die Entscheidungsgewalt über diese normativ bedeutsamen Festlegungen liegt dabei häufig bei großen Technologieunternehmen, die ohne demokratische Legitimation weitreichende gesellschaftliche Entscheidungen treffen (Pasquale, 2015). Dieses Ungleichgewicht zwischen technokratischen Entscheider:innen und Betroffenen führt somit zu einem „asymmetrisch strukturierten politischen Raum, dessen Koordinaten nur für einige wenige offenliegen“, so die Kritik Bartels (2023, S. 51), der fortführt: „Verborgen bleiben die Uneindeutigkeiten, das Nichtwissen und die kontingenten Dynamiken sowie die Praktiken sozialer Konstruktion von Technik, die durch die Behauptung zahlenmäßiger Eindeutigkeit Gefahr laufen, verdeckt zu werden“ (Bartels, 2023, S. 51). Vor diesem Hintergrund wiederholen sich die Partizipationsforderungen der 1980er und 1990er Jahre, wobei in gesellschaftlichen Aushandlungsprozessen zentrale Werte der KI-Entwicklung definiert werden sollen, um einem technokratisch delegitimierten KI-Einsatz entgegenzuwirken (Danaher et al., 2017; Lee et al., 2019; Sætra, 2020).

6.3.3. Partizipation als politische Rechtfertigung einer demokratischen Technikentwicklung

Im Zuge der sozialkonstruktivistischen Wende und der Technikkritik der 1970er und 1980er Jahre entstand die Forderung, technologische Entscheidungen nicht allein Expert:innen zu überlassen, sondern soziale Organisationen, Betroffene und Bürger:innen aktiv zu beteiligen (Wehling, 2021). Daraus entwickelten sich partizipative Steuerungsmodelle, die auf Demokratisierung und Transparenz zielten. Ihre zentrale Rechtfertigung lag in der Annahme, dass gesellschaftliche Teilhabe zur Demokratisierung der Technologien und zu legitimeren Entscheidungen führt (Bogner, 2021b; Frahm et al., 2022; Rahm, 2023a; Stilgoe et al., 2013; Weyer, 2008). Sie fand Ausdruck beispielsweise im Rahmen der ELSI-Programme (*Ethical, Legal and Social Implications*) des *Human Genome Project* (Clayton, 2003; Hilgartner et al., 2017), die Risiken frühzeitig durch gesellschaftliche Beteiligung adressieren sollten (Bogner, 2021b; Doezema & Frahm, 2023; Heyen et al., 2024). Technologiepolitik wurde damit zu einer partizipativen Gestaltungspolitik, die technische Entwicklung als kollektiven Lernprozess begreift und nachhaltige, an gesellschaftlichen Problemlagen orientierte Innovation priorisiert. Mit dem Übergang zu einem *Demand-Pull*-Ansatz**,** der technologische Entwicklung an gesellschaftlichen Bedürfnissen orientiert, wurde Partizipation selbst zu einer Rechtfertigung für die aktive demokratische Gestaltung des technologischen Wandels (Doezema & Frahm, 2023; Gassler et al., 2006; Weyer, 2008), die der Cité des Staatsbürgers zugeordnet werden kann.

6.3.3.1. Die Chance der aktiven demokratischen Gestaltung des Einsatzes Künstlicher Intelligenz

Empirische Erkenntnisse der gegenwärtigen KI-Politik verweisen darauf, dass trotz der dargelegten deterministischen Prägung eine partizipative Technikgestaltung weiterhin als politisches Steuerungsinstrument aufgegriffen wird. Der digitale Wandel in Deutschland wird zunehmend als Gestaltungsaufgabe verstanden, die Staat, Wirtschaft und Gesellschaft gleichermaßen betrifft. In politischen Debatten wird die Notwendigkeit einer aktiven Mitgestaltung betont, um die Potenziale von KI zu nutzen sowie die Wettbewerbsfähigkeit langfristig zu sichern (Lambach & Oppermann, 2023; Marenco & Seidl, 2021). Der Einsatz von KI und der digitale Wandel werden dabei in einem grundsätzlich positiven Licht dargestellt: Ihre Gestaltung gilt als Chance und nicht als Bedrohung. Sie wird als gemeinsame Aufgabe verstanden, die eine Kooperation zwischen Staat, Wirtschaft, Wissenschaft und Bürger:innen erfordert und die Möglichkeit eröffnet, KI an gesellschaftlich-demokratischen wie auch wirtschaftlichen Bedürfnissen auszurichten (Köstler & Ossewaarde, 2022; Liebig et al., 2024; Marenco & Seidl, 2021; Nokkala, 2016). Ein Schwerpunkt dieser Demokratisierungsbestrebungen liegt neben der gesellschaftlichen Partizipation ebenso auf dem Einbezug wissenschaftlicher und technischer Expertise (Radu, 2021). So wurde etwa 2018 die Datenethikkommission eingerichtet, deren

Auftrag es ist, „ethische Maßstäbe und Leitlinien sowie konkrete Handlungsempfehlungen für den Schutz des Einzelnen, die Wahrung des gesellschaftlichen Zusammenlebens und die Sicherung und Förderung des Wohlstands im Informationszeitalter zu entwickeln“ (Datenethikkommission der Bundesregierung, 2019, S. 13). Gleichzeitig erhält die Partizipation der Bevölkerung besonderes Gewicht: Bürger:innen werden in diesem Verständnis nicht mehr als passive Nutzer:innen digitaler Technologien, sondern als aktive Mitgestalter:innen des digitalen Wandels adressiert (Nokkala, 2016; Lambach & Oppermann, 2023). Die Verantwortung des Menschen für die Gestaltbarkeit technologischer Entwicklungen wird dabei insbesondere von wissenschaftlichen und zivilgesellschaftlichen Akteur:innen in den Fokus gerückt (Schopmans, 2022), die sich somit verstärkt auf die Cité des Staatsbürgers beziehen (Kajava et al., 2025). Inklusive Beteiligungsprozesse sollen Legitimität durch eine Demokratisierung der KI-Entwicklung sichern. Daran anschließend rückt der Ausbau digitaler Kompetenzen als zentrale Voraussetzung für demokratische Teilhabe und die verantwortungsvolle Gestaltung der KI-Entwicklung in den Fokus der politischen Debatte.

6.3.3.2. Investitionen in Bildung und Forschung zur Sicherung wirtschaftlicher und gesellschaftlicher Teilhabe

Der Bildungsanspruch umfasst sowohl die breite Bevölkerung, die im Sinne der staatsbürgerlichen Cité zu mündigen Bürger:innen befähigt werden soll, als auch die Qualifizierung von Fachkräften, die den Übergang zu einer „Industrie 4.0“ ermöglichen. Im Mittelpunkt steht die Verbindung technischer Kompetenzen mit einem Verständnis für den Einsatz von KI. Bildung gilt dabei in der politischen Debatte als zentrale Voraussetzung wirtschaftlicher und gesellschaftlicher Teilhabe im digitalen Zeitalter (Lambach & Oppermann, 2023; Lemke et al., 2024; Marenco & Seidl, 2021; Nemorin et al., 2023; Ossewaarde & Gülenç, 2020; Thelen, 2019). Forderungen nach einem Recht auf lebenslanges Lernen sollen Beschäftigte kontinuierlich auf die Herausforderungen der digitalen Arbeitswelt vorbereiten (Eynon & Young, 2021; Marenco & Seidl, 2021; Schopmans, 2022).

Zugleich sollen Bildungsinitiativen eng mit der Forschungsförderung sowie dem Wirtschaftssektor verknüpft werden, um Synergien zu schaffen und den Übergang zu einer digitalisierten Wirtschaft zu erleichtern (Liebig et al., 2024; Thelen, 2019). Forschung und Entwicklung gelten aus Sicht politischer Entscheider:innen als entscheidend für eine erfolgreiche Implementierung von KI. Investitionen in diese Bereiche sollen Innovationen fördern und die deutsche Wirtschaft an die Spitze der globalen Wissensökonomie führen (Thelen, 2019). Somit knüpft die politische Rahmung des digitalen Wandels an dieser Stelle ebenso an Rechtfertigungsmuster an, die der Cité der Industrie und des Marktes zugeschrieben werden können. In diesem Kontext wird

ebenfalls die Rolle von Expert:innen betont. So hebt etwa die nationale KI-Strategie die Bedeutung des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI) sowie internationaler Kooperationen hervor, um exzellente Forschung zu leisten, Wissen zu bündeln und von (europäischen) *Best-Practice*-Beispielen zu profitieren (Guenduez & Mettler, 2023; Radu, 2021).

6.3.4. Fallstricke einer beabsichtigten Politisierung durch partizipative Technikgestaltung

Während die vorliegenden empirischen Erkenntnisse der politischen Digitalpolitikdebatte darauf hinweisen, dass Forderungen einer partizipativen Technikgestaltung anerkannt werden, eröffnet die Erkenntnis, dass Technik in sozialen Kontexten ausgehandelt und mit Bedeutungen versehen wird, zugleich neue Formen der Politisierung (Hay, 2007): Technikgestaltung wird als aktiver Aushandlungsprozess verstanden, in dem unterschiedliche gesellschaftliche Gruppen konkurrierende Deutungen und normative Ansprüche einbringen (Katzenbach, 2018; Pinch & Bijker, 1984; Schulz-Schaeffer, 2021). Politisierung wird dabei nicht länger als bloße Reaktion auf technische Krisen verstanden, sondern wurde historisch betrachtet mit der Einführung partizipativer Technikgestaltung bewusst angestrebt (Bogner, 2021b).

In seiner praktischen Umsetzung wird dieser Ansatz in der techniksoziologischen Debatte jedoch aus zwei Gründen problematisiert: Erstens zeigt die historische Erfahrung, dass trotz des wachsenden Anspruchs partizipativer Technikgestaltung wissenschaftliche Expertise handlungsleitend blieb. So wurde mit dem Büro für Technikfolgenabschätzung (TAB) beim Deutschen Bundestag im Jahr 1990 ein explizit expert:innenbasiertes Modell geschaffen, zu einer Zeit der Hochphase der Forderung nach gesellschaftlicher Mitgestaltung und der Einbeziehung von Laien in Technikentscheidungen. Interdisziplinäre wissenschaftliche Expertise sollte eine prospektive Bewertung technologischer Potenziale und Risiken ermöglichen, blieb jedoch, wie Weyer (2008) kritisch anmerkt, „eine expertokratische Form der Austragung von Technikkontroversen“ (S. 277). Zugleich stellte sich bald Ernüchterung über die Grenzen partizipativer Verfahren ein. Diese zeigte sich einerseits in der Konfrontation zwischen Expert:innen und Laien, die mit einer Vermischung von Fakten und Werturteilen einherging, und andererseits in der Kritik an mangelnder Repräsentativität und Transparenz der Verfahren (Grunwald, 2003; Wehling, 2021; Weyer, 2008). Auch die gegenwärtige KI-Debatte sieht sich dieser Kritik ausgesetzt: Eine qualitative Expert:inneninterviewstudie beschreibt die Debatte rund um eine ethische und vertrauenswürdige KI-Entwicklung als eine vornehmlich durch politische Akteur:innen verfolgte Initiative und somit als: „top-down call in the name of public interest“ (Hirsch-Kreinsen & Krokowski, 2024b, S. 2924). Dadurch werden die praktischen Grenzen idealisierter partizipativer Gestaltungsprozesse deutlich.

Zweitens bleibt im Kontext sozialkonstruktivistischer Partizipationsforderungen der Anspruch einer demokratischen Gestaltung von KI vage. Wie bereits im Zusammenhang mit den Ethikleitlinien der KI-Entwicklung kritisiert wurde (Beckman et al., 2024; Bleher & Braun, 2023; Guihot et al., 2017; Larsson, 2020; Mittelstadt, 2019; Munn, 2023; Williams et al., 2022), bleibt unklar, wie die Demokratisierung der Technikgestaltung jenseits formaler Partizipation in der Praxis konkret umgesetzt werden soll. Stattdessen folgt die positive politische Deutung von KI, die ihre Gestaltung als Chance begreift, der Annahme, dass der Einbezug verschiedener Stakeholder:innen und insbesondere der Betroffenen per se ausreiche, um die Legitimität des KI-Einsatzes zu sichern. Dieser Ansatz kann als Versuch einer verfahrensbasierten *Ex-ante-Legitimierung* verstanden werden (siehe Abschnitt 3.2), bei der Legitimität primär durch Beteiligungsverfahren hergestellt werden soll. Das Gelingen dieser Form der Legitimation beruht dabei auf zwei zentralen Prämissen: zum einen auf der Annahme, dass KI entlang gesellschaftlicher Ansprüche frei gestaltbar ist und sich vollständig von technischen Eigenschaften der Technologie sowie den Intentionen ihrer Entwickler:innen trennen lässt (Ma, 2020; Passoth, 2008), wodurch jedoch etwa die Entstehung unintendierter Technikwirkungen nicht erklärt werden kann. Zum anderen setzt sie voraus, dass Legitimation in einem gleichberechtigten Aushandlungsprozess erfolgt, wodurch bestehende Machtverhältnisse und strukturelle Asymmetrien ausgeblendet werden. Dabei handelt es sich um eine sozialkonstruktivistische Annahme, die als nicht haltbar kritisiert wird (Gutiérrez, 2023; Law, 1990; Lazar, 2024; Star, 1990). Vor diesem Hintergrund wendet sich die techniksoziologische Debatte zunehmend einem kritisch-pragmatischen Technikverständnis zu.

6.4. Die moralische Fundierung der Technikgestaltung: Die Perspektive des kritischen Pragmatismus

Der Technopragmatismus löst die Gegenüberstellung von Technikdeterminismus und Sozialkonstruktivismus auf (Rammert, 2016). Dies liegt nicht zuletzt in der Anerkennung der begrenzten Erklärung politischer Technikgestaltung beider Perspektiven begründet. Während der Technikdeterminismus die Möglichkeit bewusster Gestaltung und normativer Reflexion unterschätzt, vernachlässigt der Sozialkonstruktivismus die materielle Wirkmacht und Stabilität technischer Strukturen sowie bestehende Machtungleichgewichte in technischen Entwicklungsprozessen. Daher gerät auch die Gestaltungsmacht der Politik in Konflikt mit diesen beiden Perspektiven: Indem der Technikdeterminismus die Autonomie der Technikentwicklung betont und der Sozialkonstruktivismus Technik vollständig als Produkt gesellschaftlicher Konstruktionen begreift, schränken beide Positionen die politische Einflussnahme auf technische Innovationen ein (Mensch & Schmidt, 2003; Ropohl, 1996). Demgegenüber begreift ein

technopragmatisches Verständnis das Verhältnis von Technik und Gesellschaft als eines der wechselseitigen Beeinflussung (Dolata & Werle, 2007; Latour, 2004, 2005, 2006; Volti & Croissant, 1988/2024), wobei Technik und Gesellschaft sich in einem dynamischen, ko-evolutionären Prozess gemeinsam weiterentwickeln (Dolata, 2008). Zugleich rücken damit Fragen der menschlichen *Agency* und des gesellschaftlichen Zusammenlebens mit Technik (wer nimmt in soziotechnischen Interaktionen welche Rolle ein?) in den Vordergrund (Anicker et al., 2024; Callon & Law, 1995; Dafoe, 2015; Rammert, 2016, 2025), wodurch zentrale Fragen politischer Steuerung aufgeworfen werden.

Vor diesem Hintergrund beschreibt Hughes' (1994) Konzept des *Technological Momentum* einen Wandel des Verhältnisses von Technik und Gesellschaft im Laufe der Zeit. In der frühen Entwicklungs- und Implementierungsphase ist Technik formbar und stark abhängig von sozialen, wirtschaftlichen und politischen Kontexten. Sie wird gestaltet, angepasst und von gesellschaftlichen Interessen bestimmt. Mit zunehmender Entwicklung und Verankerung in der Gesellschaft entwickelt sie jedoch ein eigenes „Momentum": Sie wird stabiler, weniger veränderbar und beginnt, selbst gesellschaftliche Strukturen zu prägen. Technik gewinnt also im Verlauf ihrer Entwicklung an Eigenständigkeit, ohne dabei ihre soziale Pfadabhängigkeit zu verlieren. Dieses Modell vereint damit deterministische und konstruktivistische Elemente und beschreibt das Verhältnis von Technik und Gesellschaft als dynamische Wechselwirkung.

Technik wird damit nicht nur als gestaltbar anerkannt, sondern zugleich als strukturierende und normativ wirksame Kraft, die Handlungsoptionen vorgibt und gesellschaftliche Ordnungsvorstellungen entweder stabilisiert oder infrage stellt. Dadurch entsteht die Notwendigkeit einer aktiven und öffentlich legitimierten Gestaltung. Zugleich wird in diesem pragmatischen Technikverständnis ein erhöhter Rechtfertigungsdruck politischer Entscheidungen sichtbar. Wenn Technik einerseits tief in gesellschaftliche Ordnungen eingreift und andererseits als kontingent und veränderbar erkannt wird, ist jede Entscheidung über ihre Entwicklung und ihren Einsatz potenziell strittig und muss daher legitimiert werden (Diaz-Bone & Schrör, 2022; Selk, 2011). Rechtfertigung wird damit zu einem zentralen Element politischer Entscheidungsfindung unter Bedingungen von Unsicherheit und Kontingenz.

Gerade im Fall von KI verweist das Konzept des *Technological Momentum* dabei auf den hohen Gestaltungs- und Rechtfertigungsbedarf in der aktuellen Entwicklungsphase. Einerseits ist die heutige Form von KI-Anwendungen historisch betrachtet eine noch junge Technologie, andererseits zeigen ihre zunehmende Verbreitung in Verwaltung, Bildung und anderen gesellschaftlichen Bereichen sowie ihre ambivalenten Einflussmöglichkeiten die Dringlichkeit eines

Aushandlungsprozesses, der nicht nur technisches Fachwissen, sondern ebenso normative Erwartungen berücksichtigt. Die Gegenüberstellung optimistischer und pessimistischer Zukunftsszenarien verdeutlicht dabei die Unsicherheit über die tatsächliche Entwicklungsrichtung (siehe Abschnitt 6.3.1). Vor diesem Hintergrund rückt die Frage in den Mittelpunkt, welche moralischen Werte die Gestaltung und Nutzung von KI leiten:

> As the adoption of AI by governments to enhance efficiency in service provision may result in supplanting or otherwise reprioritizing key public values, there is a need to better understand which public values should guide government implementation of AI systems. This is especially true at the present moment when major implementation, procurement, and regulatory decisions are being made that will shape AI's long-term impacts in the public domain. (Schiff et al., 2022, S. 669)

Die Aushandlung der zugrunde liegenden Werte der KI-Politik ist dabei im Kontext einer zunehmend netzwerkbasierten Innovationspolitik zu verstehen. Technikentwicklung wird dabei als gemeinsames Projekt staatlicher, wissenschaftlicher, wirtschaftlicher und gesellschaftlicher Akteur:innen verstanden. Dabei sollen die Verbindung wissenschaftlicher Expertise und partizipativer Verfahren dazu beitragen, moralische Grundlagen auszuhandeln, potenzielle Risiken und Konflikte frühzeitig zu erkennen und gesellschaftliche Akzeptanz zu sichern (Bogner, 2021b; Heyen et al., 2024). Die Bedeutung der moralischen Grundlage der Technikgestaltung kann dabei anhand des Zugangs der Soziologie der Quantifizierung nachvollzogen werden. Sie zeigt, dass Datenerhebung, Klassifikation und Bewertung keine rein technischen Routinen darstellen, sondern Ausdruck kollektiver Aushandlungsprozesse über gesellschaftliche Werte und Normen sind. Aus dieser Perspektive lässt sich nachvollziehen, wie sich politische Zielvorstellungen und soziale Konventionen in den technischen Strukturen von KI-Systemen materialisieren und wie dadurch Legitimität erzeugt oder infrage gestellt werden kann.

6.4.1. Das pragmatische Verständnis der moralischen Konstruktion Künstlicher Intelligenz

Die durch Desrosières (1998) vorgelegte Soziologie der Quantifizierung[16] knüpft an ein pragmatisches Verständnis sowohl der Technikgestaltung als auch des Legitimationsprozesses an. Es folgt den Grundannahmen der SdK und betont die Rolle moralischer Grundsätze in den sozialen Aushandlungsprozessen technischer Gestaltung. Im Zentrum steht die Einsicht, dass sowohl die Messbarmachung sozialer Phänomene als auch ihre algorithmische Verarbeitung in KI-Systemen von moralischen Überzeugungen beeinflusst sind (Desrosières, 1998, 2001; Diaz-

[16] Die Soziologie der Quantifizierung ist originär von Desrosières nicht im KI-Kontext gedacht, bietet jedoch aufgrund der Verkörperung des Quantifizierungsstrebens durch Big Data und KI starke Analogien (vgl. z. B. Henneguelle, 2023; Kappler, 2025).

Bone et al., 2020; Kappler, 2025). Die Entwicklung und Anwendung von KI beruht auf einer Vielzahl von Entscheidungen, die von Entwickler:innen, Anwender:innen und politischen Entscheidungsträger:innen getroffen werden und stets von individuellen wie kollektiven Konventionen geleitet sind. Damit wird Technikgestaltung als sozialer Aushandlungsprozess sichtbar: von der Auswahl und Definition der Daten über die Modellarchitektur bis zur Interpretation der Ergebnisse (Henneguelle, 2023; Ma, 2020; Weinberg, 2022). Desrosières (1998) hebt hervor, dass jeder Form der Quantifizierung soziale Übereinkünfte über Kategorien und Definitionen vorausgehen. In Analogie dazu erfordert die Entwicklung von KI-Systemen normative Festlegungen darüber, welche Daten als relevant gelten, welche Merkmale extrahiert werden und wie Ergebnisse bewertet werden (Henneguelle, 2023; Salais, 2016). Die Entscheidung, wie etwa im Bildungsbereich im Kontext einer KI-basierten Leistungsprognose zu definieren, wer als „studienabbruchsgefährdet" gilt, bestimmt unmittelbar die Auswahl der Trainingsdaten und wirkt sich auf die gesamte Modellstruktur aus. Die Datengrundlage von KI ist somit kein objektives Abbild der Realität, sondern Ausdruck gesellschaftlicher Konventionen (Cités):

> every quantification must be preconditioned by a societal ‚convention'—the social process by which its members agree on how to encode, compare, and count which objects. Therefore, statistical facts exhibit social constructivity. They are not objective responses to pure facts but determined by temporary social conventions, which can be altered with a change of indicators and parameters. (Li, 2024, S. 2)

Auch die Wahl von Algorithmen und Modellen erfolgt innerhalb bestehender methodischer Standards und institutioneller Praktiken. Entscheidungen darüber, wie neuronale Netze trainiert werden, welche Gewichtungen oder Validierungskriterien gelten, beruhen auf sozialen und professionellen Übereinkünften. Selbst mathematisch-technische Prozesse sind daher normativ durchdrungen (Henneguelle, 2023). Ebenso ist die Interpretation der Ergebnisse abhängig von moralischen Ansprüchen, da KI-basierte Klassifikationen und Prognosen immer auf vordefinierten Kategorien basieren, deren Bedeutung sozial ausgehandelt wird. Fragen wie „Was gilt als erfolgreiche Vorhersage?" oder „Welche Metriken sind relevant?" zeigen, dass technisches Wissen stets in gesellschaftliche Bewertungsrahmen eingebettet bleibt (Desrosières, 1998).

Insgesamt wird deutlich, dass KI-Systeme auf Entscheidungen beruhen, die durch gesellschaftlich etablierte Cités strukturiert sind. Diese prägen, was als messbar, verarbeitbar und legitim gilt, und machen sichtbar, dass KI-Entwicklung nicht als rein technischer, sondern als sozio-technischer Prozess zu verstehen ist. Die Soziologie der Quantifizierung relativiert damit den Anspruch technischer Objektivität zugunsten einer Technikdeutung, die kontingent und

interpretativ verhandelbar bleibt (Thévenot, 2016). Durch die Offenlegung der moralischen und politischen Grundlagen der KI-Entwicklung eröffnet diese pragmatische Perspektive zugleich Raum für neue Formen der Kritik und Rechtfertigung von KI sowie für ihre Politisierungsprozesse.

6.4.2. Technik als Träger von Moral und Machtverhältnissen und die Notwendigkeit der Kritik

Mit der pragmatischen Auffassung, dass Technik im Laufe ihrer Stabilisierung zunehmend eigenständige Wirkmacht entfaltet, wird ihre normative und politische Dimension hervorgehoben. Wenn Technik beginnt, soziale Ordnungen mitzugestalten, wird zugleich sichtbar, dass in ihre Strukturen nicht nur moralische Entscheidungen, sondern auch Machtverhältnisse eingeschrieben sind (Crawford, 2021; Kalluri, 2020). Besonders KI-Systeme tendieren durch ihre Orientierung an Standardisierung und Hierarchisierung dazu, bestehende Ungleichheiten zu verstärken (Henneguelle, 2023; Ma, 2020; Weinberg, 2022). Kritik an diesen Machtasymmetrien zeigt sich zunehmend im wissenschaftlichen Diskurs, so etwa im Bildungsbereich, wo Lernende sich den eingesetzten Systemen kaum entziehen können, während Bildungseinrichtungen und Technologieanbieter umfangreiche Daten verarbeiten und auf deren Grundlage folgenreiche Entscheidungen treffen (Castañeda & Selwyn, 2018; Giró Gràcia & Sancho-Gil, 2021; Macgilchrist, 2019; McConvey & Guha, 2024). Damit richtet sich der Blick auf die Frage, welche Werte und Prioritäten (z. B. Effizienz, Kontrolle oder Standardisierung) in technische Systeme eingeschrieben werden und wessen Interessen sie repräsentieren (Feenberg, 1991, 1999, 2009; Jasanoff, 2004, 2016).

Feenberg (1999, 2002) fordert daher im Zuge eines kritisch-pragmatischen Ansatzes die Berücksichtigung der moralischen Grundlagen der Technikentwicklung und ihrer politischen Regulierung, um die Einbettung der Technik in ihren sozialen Kontext umfassend untersuchen zu können. Technikgestaltung wird hier als ein Prozess verstanden, der Kritik und moralische Reflexion ausdrücklich einschließt. Eine Perspektive, die bereits im pragmatischen Legitimitätsverständnis zum Ausdruck kommt (siehe Abschnitt 3.3). So wird am Beispiel der Kritik der 68er-Bewegung an der technokratischen Durchdringung gesellschaftlicher Bereiche der Einfluss auf Technik durch ihre kritische Infragestellung sichtbar: „With a public questioning technology, demanding changes and in some cases carrying them out, technology is thus more flexible, transformable and amenable to democratic debate and reconstruction“ (Kellner, 2001, S. 158). Technologische Legitimation entsteht hier nicht allein aus Effizienz und Funktionalität oder idealisierten Partizipationsprozessen, sondern durch kritische Aushandlungsprozesse und moralische Reflexion, die Technik demokratisch rückbinden.

Auf dieser Grundlage gewinnt die kritische Perspektive besondere Bedeutung für die gegenwärtige KI-Entwicklung. Sie ermöglicht es, dominante Technikdeutungen und Machtstrukturen, die die Integration von KI prägen, zu hinterfragen (Couldry, 2020; Selwyn, 2024). Im Bildungsbereich etwa kann Kritik Reflexivität beispielsweise über Fragen der Ungleichheit fördern und zugleich alternative Visionen eröffnen, in denen moralische Grundsätze Vorrang vor marktwirtschaftlichen oder solutionistischen Zielen erhalten. Ohne diese kritische Auseinandersetzung droht der KI-Einsatz, bestehende Ungleichheiten zu verfestigen und demokratische Ziele zu untergraben (Macgilchrist, 2021; Selwyn, 2024).

Das wachsende Bewusstsein über die moralische Fundierung technischer Entwicklung und ihre Bedeutung für politische Legitimation scheint sich (zumindest oberflächlich) auch in der Ausrichtung der Technologiepolitik widerzuspiegeln, wie nachfolgend auf Basis der empirischen Erkenntnisse einer wertebasierten KI-Politik dargelegt wird. Im Anschluss wird die techniksoziologische Kritik der gegenwärtigen KI-Debatte dargelegt.

6.4.3. Die politische Rechtfertigung einer wertebasierten Entwicklung Künstlicher Intelligenz

Historisch betrachtet verzeichnete sich im Zuge des wachsenden Bewusstseins für die moralische Fundierung technischer Entwicklung ein „*normative turn*“ der Innovationspolitik, der technologische Fortschritte nicht mehr allein an ökonomischen, sondern zunehmend an demokratischen Werten ausrichtet (Uyarra et al., 2019). Programme wie das von der EU (2014) verfolgte Konzept der *Responsible Research and Innovation* (Owen et al., 2013; Owen & Pansera, 2019; Stilgoe et al., 2013) sollen sicherstellen, dass technologische Entwicklungen sozialverantwortlich gestaltet werden. Ziel ist es, Risiken und unbeabsichtigte Folgen frühzeitig zu erkennen und Entscheidungen am Gemeinwohl zu orientieren (Bogner, 2021b; Heyen et al., 2024; Rahm, 2023a; Uyarra et al., 2019).

Die empirischen Erkenntnisse über die politische KI-Debatte belegen, dass die Gestaltung einer menschenzentrierten KI-Entwicklung, die zentrale gesellschaftliche Werte wie soziale Teilhabe, individuelle Freiheit, Sicherheit und Nachhaltigkeit berücksichtigt, zunehmend an Bedeutung gewinnt (Bareis & Katzenbach, 2022; Guenduez & Mettler, 2023; Köstler & Ossewaarde, 2022; Ossewaarde & Gülenç, 2020; von Essen & Ossewaarde, 2023). Diese Bestrebungen können als Ansprüche der Cité des Staatsbürgers und der Ökologie verstanden werden. Auch auf EU-Ebene zielt die KI-Regulierung auf eine an europäischen Werten orientierte Entwicklung, die sowohl die Legitimität politischer Steuerung erhöhen als auch die europäische Integration stärken soll (Frahm & Schiølin, 2023; von Essen & Ossewaarde, 2023). Dabei zeigt sich im Kontext einer Analyse skandinavischer Technologiepolitik, dass die politische Steuerung von KI pfadabhängig und stark von nationalen Wertvorstellungen geprägt ist (Robinson,

2020). In Deutschland zeigen vereinzelte Analysen, dass damit etwa ethische Fragen des Datenschutzes sowie fairer Arbeitsbedingungen in den Fokus rücken, wobei unterschiedliche Akteursgruppen je unterschiedliche Schwerpunktthemen setzen (Lemke et al., 2024). Politische und insbesondere zivilgesellschaftliche Akteur:innen, wie Gewerkschaften, betonen die soziale Verantwortung in der digitalen Transformation der Arbeitswelt, die neben Gerechtigkeitsfragen auch den Schutz vor automatisierter Überwachung und die Gewährleistung menschenwürdiger Arbeitsverhältnisse umfasst (Marenco & Seidl, 2021). Ebenso gilt der Datenschutz als zentrales Gut zur Wahrung der Persönlichkeitsrechte. Zugleich eröffnet dieser Anspruch jedoch ein zentrales Spannungsfeld zwischen der Notwendigkeit umfangreicher Datensammlungen zur Sicherung von Innovation und Wettbewerbsfähigkeit und Bedenken hinsichtlich des Schutzes der Privatsphäre oder des Datenmissbrauchs (Guenduez & Mettler, 2023; Hirsch-Kreinsen & Krokowski, 2024b; Lambach & Oppermann, 2023; Willems, 2021).

Dieses Spannungsfeld zwischen Innovationsstreben und werteorientierter Technikregulierung verweist somit auf eine zentrale Herausforderung politischer Techniksteuerung. Zugleich wird eine diagnostizierte Priorisierung des Innovationsstrebens gegenüber einer fundierten moralischen Aushandlung zunehmend Gegenstand techniksoziologischer Kritik.

6.4.4. Zwischen Wertorientierung und technokratischer Verkürzung gesellschaftlicher Herausforderungen: Die techniksoziologische Kritik des Solutionismus

Dabei richtet sich die techniksoziologische Kritik gegen die oftmals diffuse Verwendung von Werten in politischen und ethischen Leitlinien, die den Eindruck normativer Eindeutigkeit erzeugen, ohne die strukturellen Spannungen und sozialen Machtverhältnisse der KI-Entwicklung zu reflektieren. Insbesondere im Rahmen der Solutionismuskritik wird problematisiert, dass komplexe soziale Probleme technokratisch verkürzt und als Mittel zur Legitimation technologischer Lösungen instrumentalisiert werden. Statt die moralische und politische Komplexität von KI zu berücksichtigen, tendieren solche Ansätze dazu, gesellschaftliche Probleme in technische Optimierungsaufgaben zu übersetzen, wie im Folgenden kritisch beleuchtet wird.

6.4.4.1. Die Kritik an der Verkürzung der moralischen Gestaltung Künstlicher Intelligenz

Trotz der Forderungen einer menschenzentrierten KI-Entwicklung und der Betonung europäischer Werte scheinen solutionistische, industrielle und marktorientierte Rechtfertigungen die deutsche Digital- und KI-Politik zu leiten. Ethische und soziale Implikationen der KI sowie unintendierte Risiken scheinen hingegen angesichts deterministischer Annahmen und der verbreiteten Utopie des KI-Einsatzes in der politischen Debatte über den digitalen Wandel und den Einsatz von KI nachrangig behandelt zu werden (Köstler & Ossewaarde, 2022; Lambach & Oppermann, 2023; Ossewaarde & Gülenç, 2020). Werden technologische Risiken thematisiert,

so werden diese nicht selten mit dem Verweis auf Expert:innenlösungen und anhaltende Forschung als überschaubar und kontrollierbar dargestellt, wie internationale Untersuchungen des politischen Diskurses belegen (Guenduez & Mettler, 2023; Smallman, 2018; Ulnicane & Aden, 2023). Ethische Leitlinien dienen dabei oftmals als reaktiver Problemlösungsansatz für öffentlich artikulierte KI-Kritik (Ulnicane et al., 2021, 2022). So wollen diese Legitimität erzeugen und bestehende Kontroversen entschärfen.

Gleichzeitig bleibt die Vision einer an nationalen Werten ausgerichteten KI-Entwicklung vage und mitunter inkonsistent (Bareis & Katzenbach, 2022; Köstler & Ossewaarde, 2022; Rahm, 2023a). Während in politischen Programmen wiederholt auf die zentrale Bedeutung einer wertebasierten KI für die Zukunft Deutschlands verwiesen wird, mangelt es hingegen häufig an konkreten Vorstellungen darüber, wie diese Zukunft gestaltet werden soll: „While the German government's frame of *AI as key to the German future* blurs the meaning of AI, its frame of *German AI* blurs the role of German values in shaping the German AI future" (Köstler & Ossewaarde, 2022, S. 256; Herv. i. O.). Diese Diffusität in Bezug auf das Verständnis von KI und ihrer moralischen Grundlagen lässt sich dabei nicht nur im politischen Diskurs, sondern ebenfalls im Expert:innen-KI-Diskurs verzeichnen (Schopmans, 2022). Der zugeschriebene Einfluss der Technologie auf die Gesellschaft variiert etwa in Abhängigkeit davon, ob KI im Allgemeinen als Schlüsseltechnologie der Zukunft aufgefasst wird oder ob spezifische KI-Anwendungen diskutiert werden. Dieses uneindeutige Bild der KI steht gewissermaßen im Widerspruch zur skizzierten (deterministischen) deutschen Zukunft, die ohne KI nicht denkbar sei und der ein großes Potenzial für die Gesellschaft und Wirtschaft zugesprochen wird (Köstler & Ossewaarde, 2022; siehe Abschnitt 6.1.1). Dabei wird KI nicht nur als unausweichlich dargestellt, sondern zugleich als Technologie inszeniert, die alle bisherigen Innovationen übertrifft und damit frühere technologische Wandlungsprozesse in den Schatten stellt. Solche Zuschreibungen gehen nicht nur mit einer Verklärung technologischer Kontinuitäten und einer Ausblendung historischer Erfahrungen einher, sondern stellen auch eine Form der Komplexitätsreduktion dar, die der Depolitisierung dient. Indem der Handlungsspielraum auf die Frage verengt wird, *wie* KI eingesetzt werden soll, ohne zu hinterfragen, *ob* sie ein angemessenes Mittel zur Bearbeitung gesellschaftlicher Probleme darstellt, wird politische Gestaltbarkeit unterminiert (Bareis & Katzenbach, 2022; Mosco, 2004).

Diese Beobachtung verweist auf ein zentrales Muster gegenwärtiger Technikkritik: die Neigung, komplexe gesellschaftliche Herausforderungen durch technologische Lösungen bewältigen zu wollen. Diese solutionistische Denkweise begreift KI als inhärent fortschrittlich (siehe Abschnitt 6.1.1.1) und lässt soziale wie politische Aushandlungsprozesse in den Hintergrund

treten. Im folgenden Abschnitt wird diese Solutionismus-Kritik genauer beleuchtet, die den Glauben an KI als universelles Problemlösungsinstrument als eine Quelle gegenwärtiger Legitimationslogiken hinterfragt.

6.4.4.2. Unsachgemäße Komplexitätsreduktion: Die Kritik der solutionistischen Rechtfertigung

Solutionistisch-deterministische Rechtfertigungen von KI beruhen auf der Annahme, gesellschaftliche Probleme ließen sich durch technische Innovationen optimal lösen. Diese Technikgläubigkeit zeigt sich nicht nur in der stetigen Ausweitung von KI-Anwendungsfeldern, sondern auch in Phänomenen wie dem sogenannten *„Function Creep“* und dem *„Second-Order-Solutionism“*. Ersteres beschreibt eine schleichende Ausweitung ursprünglich zweckgebundener Anwendungen auf neue Einsatzbereiche, häufig ohne angemessene gesellschaftliche oder rechtliche Kontrolle (Grimm et al., 2021; Koops, 2021; Ulbricht, 2020). Zweiteres bezeichnet die Tendenz, technische Antworten auf neue technisch induzierte Probleme zu geben, wie sie beispielsweise bereits von Postman (1992) oder Luhmann (1991) vor geraumer Zeit thematisiert wurden. Auf diese Weise entsteht ein „second-order solutionism where they promise technological solutions to problems that their own technologies have created“ (Nachtwey & Seidl, 2023, S. 17).

Kritisiert wird zudem die wachsende Forderung nach technischem Kompetenzaufbau als Voraussetzung gesellschaftlicher Zukunftsfähigkeit. Im britischen Diskurs etwa gilt die Fähigkeit zu programmieren als unverzichtbar für die Teilhabe an der digitalen Gesellschaft. Williamson (2016) sieht in dieser Forderung des „Learning to Code“ jedoch selbst einen Ausdruck des Solutionismus, da die Beherrschung digitaler Werkzeuge beziehungsweise verbreitete Programmierfähigkeiten als zentrale Anforderung vermittelt wird, viele der aktuellen wirtschaftlichen und gesellschaftlichen Probleme zu lösen. Diese Denkweise verengt komplexe gesellschaftliche Herausforderungen auf technische Optimierungsfragen und blendet normative Dimensionen aus (Smithers, 2020). Indem technologische Lösungen als alternativlos gelten, werden politische Aushandlung und moralische Reflexion verdrängt.

Die Kritik des Solutionismus richtet sich daher gegen diese Verkürzung sozialer Komplexität und die Ignoranz gegenüber alternativen, nichttechnischen Formen der Problembearbeitung. Morozov (2013) zufolge werden auf Basis solutionistischer Ansätze vorschnell technologische Antworten formuliert, ohne die Problemdefinition zuvor kritisch zu hinterfragen:

> [S]olutionism presumes rather than investigates the problems that it is trying to solve, reaching ‚for the answer before the questions have been fully asked.’ How problems are composed matters every bit as much as how problems are resolved.

> Solutionism, thus, is not just a fancy way of saying that for someone with a hammer, everything looks like a nail; it's not just another riff on the inapplicability of ‚technological fixes' to ‚wicked problems' (…). It's not only that many problems are not suited to the quick-and-easy solutionist tool kit. It's also that what many solutionists presume to be ‚problems' in need of solving are not problems at all; a deeper investigation into the very nature of these ‚problems' would reveal that the inefficiency, ambiguity, and opacity–whether in politics or everyday life–that the newly empowered geeks and solutionists are rallying against are not in any sense problematic. (Morozov, 2013, S. 6)

Im Fokus der Kritik solutionistischer Denkweisen steht somit insbesondere eine verzerrte Problempriorisierung und -definition. Statt komplexe gesellschaftliche Herausforderungen in ihrer Tiefe zu analysieren, greifen solutionistische Ansätze häufig vorschnell zu technologischen Antworten auf oberflächlich definierte Probleme. Dadurch bevorzugt der Solutionismus einerseits die Bearbeitung bestimmter gesellschaftlicher Probleme gegenüber anderen – nämlich eben jenen, die eine technische Antwort erlauben (Genus & Stirling, 2018; Khanal et al., 2025; Morozov, 2013; Stilgoe, 2020). Andererseits institutionalisiert er damit zugleich spezifische Wertvorstellungen, indem einfacher zu messende soziale Phänomene gegenüber komplexen, schwer quantifizierbaren Problemen priorisiert werden (Mau, 2020). Gleichzeitig werden nichttechnische Problemlösungen ausgeschlossen (Green & Viljoen, 2020; Weinberg, 2022; Williamson, 2016). Diese unsachgemäße Komplexitätsreduktion sozialer Herausforderungen auf technische Funktionalitäten wird insbesondere in sensiblen Bereichen wie dem Bildungswesen kritisiert (Giró Gràcia & Sancho-Gil, 2021; Horvath et al., 2025; Rahm & Rahm-Skågeby, 2023; Sharma, 2022).

Auf diese Weise wird nicht nur der gesellschaftliche Aushandlungsprozess über die Frage, *was* überhaupt als Problem gilt, umgangen, sondern auch der Blick auf strukturelle soziale Ursachen verstellt. Im Zentrum stehen symptomorientierte Interventionen, die kurzfristige Effizienz versprechen, jedoch grundlegende soziale Zusammenhänge unberührt lassen. So wird die technische Lösung zum Ersatz für eine tiefgreifende Ursachenanalyse. Dies hat zur Folge, dass die eigentlichen Problemlagen nicht nachhaltig bearbeitet, sondern lediglich funktional verwaltet werden (Broussard, 2018; Morozov, 2013, 2014; Stilgoe, 2020; Yeung, 2018).

Darüber hinaus führt die solutionistische Rechtfertigung von KI nicht nur zu einer Verkürzung sozialer Komplexität, sondern auch zu einer Verengung der Erkenntnisse der Wissensproduktion. Wie bereits aufgezeigt wurde, stützt sich das solutionistische Versprechen auf die epistemische Autorität von Expert:innen zur Lösung politischer Probleme (siehe Abschnitt 6.1.1.1).

Wie Bartl (2023) jedoch kritisch anmerkt, ist der Rückgriff auf wissenschaftliches Wissen in solchen Kontexten häufig selektiv: Politisch anschlussfähig sind vor allem jene Disziplinen, die quantifizierende, scheinbar objektive und effiziente Lösungen anbieten. Diese Praxis verkenne jedoch nicht nur die „Brüchigkeit des Wissens" (S. 46), indem sie die Dynamik wissenschaftlicher Erkenntnis zwischen Uneindeutigkeit und Nichtwissen ignoriere, sondern blende auch die soziale Konstruiertheit und Wertgebundenheit technologischer Lösungen systematisch aus. Problematisch sei dies insbesondere dann, wenn die Fehleranfälligkeit von KI und ihre unintendierten Effekte unberücksichtigt blieben (Veale & Brass, 2019). Daher führen die Ideologie des „technological fix" und des Solutionismus Bartl (2023) zufolge

> zu einer Unterschätzung der disruptiven Potenziale neuer Technologien und einer folgenreichen Verwechslung von Daten und Wissen sowie zwischen Korrelation und Kausalität, insbesondere im Kontext von Big Data-Analysen. Aus der Sammlung, Analyse und Bewertung von Daten sowie der Externalisierung von politischer Verantwortung auf privatwirtschaftliche Akteure ergeben sich umfangreiche Implikationen, die in einem Demokratiedefizit kulminieren. Die Digitalisierung des Politischen birgt insofern einige Gefahren für demokratische Prozesse und Fragen von Legitimität, die sich nicht zuletzt aus der sozialen Konstruiertheit der entsprechenden Technologien ableitet. (S. 48)

Solutionistische Vorstellungen schließen damit direkt an ein technokratisches Bewusstsein an: Sie binden zwar technisch-wissenschaftliche Expertise in politische Entscheidungsprozesse ein, entpolitisieren diese jedoch zugleich, indem gesellschaftliche Probleme nicht mehr als genuin politische Aushandlungsfragen verstanden und politisch-demokratische Entscheidungsbefugnisse durch technologische Antworten substituiert werden (Nachtwey & Seidl, 2021).

6.4.5. Die Forderung der Politisierung Künstlicher Intelligenz aus technopragmatischer Perspektive

Diese technokratische Verschiebung im Umgang mit Technologie trägt dazu bei, dass gesellschaftliche Auseinandersetzungen über ihre politischen, sozialen und normativen Grundlagen an Intensität verlieren. Während frühere Technikkritiken, etwa im Kontext der Atomkraft oder Gentechnik, noch ausgeprägte öffentliche Kontroversen und Protestbewegungen hervorbrachten (wenngleich nicht selten erst infolge technischer Fehlschläge), scheinen diese kritischen Auseinandersetzungen in den letzten Jahren einer Phase relativer Konfliktfreiheit gewichen zu sein (Heyen et al., 2024). Trotz historischer Erfahrungen mit technologischen Risiken prägen seit Beginn der 2000er Jahre zunehmend „naiv anmutende Visionen technologischer Weltverbesserung" (Bogner, 2021b, S. 51) das öffentliche Bild technologischer Entwicklung.

Vor dem Hintergrund des deterministischen Fortschrittsversprechens und der bislang vorliegenden Erkenntnis eines weitestgehend optimistischen, gar utopischen politischen KI-Diskurses (Bareis & Katzenbach, 2022; Köstler & Ossewaarde, 2022; Ossewaarde & Gülenç, 2020) erscheint die gegenwärtige KI-Politisierung somit gering (Dumbach et al., 2023; König & Wenzelburger, 2019; Siewert & König, 2019). Gleichzeitig verdeutlicht der von Ulbricht (2020) skizzierte Politisierungszyklus, dass die Legitimität des KI-Einsatzes nicht allein von seiner technischen Leistungsfähigkeit abhängt, sondern wesentlich dadurch beeinflusst wird, ob ihre normativen Grundlagen rechtzeitig öffentlich verhandelt werden. Dabei prägt die Art und Weise, wie KI in der politischen Debatte dargestellt wird, maßgeblich, welche politischen Handlungsoptionen in Betracht gezogen werden (Bareis & Katzenbach, 2022; Köstler & Ossewaarde, 2022; Nokkala, 2016; Taeihagh, 2021). Beispielsweise kann die Problematisierung von Fragen des Datenschutzes, der Fairness oder der Arbeitsbedingungen einen wichtigen Hinweis für einen regulatorischen Handlungsbedarf liefern (Marenco & Seidl, 2021; Thelen, 2018).

Eine frühzeitige Politisierung ist entscheidend, um eine rein reaktive politische Krisensteuerung zu vermeiden, wie sie in der historischen Betrachtung der Technologiepolitik sichtbar geworden ist (Bogner, 2021b). Sie kann dazu beitragen, spätere Konflikte zu verhindern, demokratische Kontrolle und Legitimität zu sichern und die gesellschaftliche Verantwortung im Umgang mit KI zu stärken. Statt einer unkritischen Befolgung des KI-Hypes bedarf es eines reflektierten Aushandlungsprozesses der KI-Einführung, der ihrer Komplexität gerecht wird. Dabei gilt es, Risiken, unbeabsichtigte Folgen und ethische Fragen zu berücksichtigen, anstatt Innovation per se als wünschenswert zu behandeln (Heyen et al., 2024; Schlaile et al., 2017). Eine solche Legitimation erfordert die Verbindung von Kritik und moralischer Fundierung des Technikeinsatzes (Diaz-Bone & Schrör, 2022; Feenberg, 1999).

Vor diesem Hintergrund fordern verschiedene Autor:innen eine systematische Auseinandersetzung mit den normativen Grundlagen der KI-Politik und der Frage, welche Werte die staatliche KI-Implementierung und -Regulierung leiten (sollen) (R. Boyd & Holton, 2018; Mittelstadt et al., 2016; Schiff et al., 2022; Tallberg et al., 2023). Dieser Bedarf wird nicht zuletzt vor dem Hintergrund der demokratietheoretischen Herausforderung durch die Verbreitung von KI-Systemen sichtbar: Wenn KI …

> Interpretationen von Normalität und Abweichung produziert, die immer tiefer in gesellschaftliche Prozesse und Interaktionen einsickern und sich verselbständigen, spricht das für die Etablierung einer subtilen Form der Macht und Herrschaft, deren Legitimität weitestgehend unhinterfragt hingenommen wird. Umso wichtiger erscheint vor diesem

> Hintergrund die Beschäftigung mit ethischen Kriterien, und zwar sowohl im Entwicklungs- wie im Verwendungskontext von digitalen Technologien. (Bartl, 2023, S. 55)

Die Aushandlung, ebenso wie die Untersuchung der moralischen Grundlagen der Legitimation von KI, wird damit zu einer zentralen Aufgabe gegenwärtiger Forschung und Politikgestaltung. Sie muss klären, wie normative Maßstäbe in die Gestaltung, Regulierung und Rechtfertigung von KI-Systemen eingehen. Das folgende Kapitel widmet sich daher den Lehren und Desideraten einer legitimen KI-Gestaltung und zeigt auf, wie moralische und politische Rechtfertigungen zusammenspielen, um die Legitimität technologischer Innovationen in der digitalen Gesellschaft zu sichern.

7. Lehren und Desiderate der techniksoziologischen und technologie-politischen Auseinandersetzung für die Legitimation Künstlicher Intelligenz

Nachdem die techniksoziologischen Debatten sowie die bislang gewonnenen empirischen Erkenntnisse zur Kritik und Rechtfertigung des KI-Einsatzes und früherer technischer Innovationen sowie deren Konsequenzen für die Politisierung dargestellt wurden, werden nun Lehren für die weitere Forschung gezogen. Auf Basis der empirischen Erkenntnisse sowie der theoretischen Perspektiven der SdK und der STS werden Desiderate identifiziert und zentrale Forschungsfragen abgeleitet. Diese adressieren die Art und Weise, auf Basis welcher moralischen Grundlage KI in politischen Debatten gerechtfertigt und kritisiert wird, welche Konfliktlinien sich dabei abzeichnen und inwiefern deterministisch-solutionistische Grundannahmen in der politischen Debatte sichtbar werden. Auf Basis der bisherigen theoretischen und empirischen Analyse werden hierzu Annahmen darüber entwickelt, welche moralischen Rechtfertigungs- und Kritikmuster in der bildungspolitischen KI-Debatte plausibel erscheinen. Diese Annahmen werden ausdrücklich nicht in Form statistisch überprüfbarer Hypothesen formuliert, sondern dienen im Sinne eines explorativen Erkenntnisinteresses als heuristische Orientierungspunkte für die anschließende empirische Analyse. Dies erscheint insbesondere daher geboten, da es sich bei der politischen KI-Debatte um ein hochdynamisches und in weiten Teilen empirisch noch unzureichend erschlossenes Feld handelt.

In den vergangenen Jahren ist KI zunehmend zum Gegenstand öffentlicher und politischer Auseinandersetzungen geworden. Trotz ihrer wachsenden gesellschaftlichen und politischen Relevanz liegen bislang jedoch nur wenige systematische empirische Studien zum politischen KI-Diskurs in Deutschland vor. Dies steht im Kontrast zur vergleichsweise gut erforschten medialen Berichterstattung über KI (z. B. Brennen et al., 2022; Chuan et al., 2019; Fast & Horvitz, 2017; Fischer & Puschmann, 2021; Köstler & Ossewaarde, 2022; Ouchchy et al., 2020;

Vergeer, 2020; Zeng et al., 2022) sowie zur umfangreichen Forschung zu Fragen der Legitimitäts-, Akzeptanz- und Fairnesswahrnehmung algorithmischer Entscheidungssysteme selbst (z. B. Lünich & Keller, 2024; Marcinkowski et al., 2020; Martin & Waldman, 2022; Shin, 2021; Starke & Lünich, 2020; Wenzelburger et al., 2022). Die vorgelagerte Frage, wie der grundsätzliche Einsatz von KI politisch legitimiert wird, ist bislang nur unzureichend beleuchtet worden.

Die in dieser Arbeit vorausgegangene Einordnung in historische und aktuelle Bezüge eines techniksoziologischen Analyserahmens zu Fragen von Mustern der Kritik und Rechtfertigung soziotechnischer Systeme sowie ihre Verknüpfung mit empirischen Erkenntnissen der gegenwärtigen Technologiepolitik erlaubt es dennoch, einige Lehren zu ziehen. So konnte erstens dargelegt werden, dass techniksoziologische Deutungsmuster darauf einwirken, wie Technologien gerechtfertigt und kritisiert werden. Sie beeinflussen damit auch, welche moralischen Grundsätze als Legitimationsgrundlage ihrer politischen Steuerung als angemessen gelten. Zweitens deuten die bislang vorliegenden empirischen Befunde auf eine dominante politische Rechtfertigung von KI als strategische Zukunftstechnologie hin, die mit Erwartungen wirtschaftlichen Wachstums, internationaler Wettbewerbsfähigkeit und gesellschaftlicher Problemlösung verknüpft wird (Bareis & Katzenbach, 2022; Köstler & Ossewaarde, 2022; Lambach & Oppermann, 2023; Ossewaarde & Gülenç, 2020; Ulnicane et al., 2022). Diese ökonomisch und solutionistisch geprägten Narrative werden zwar häufig durch Bezugnahmen auf eine partizipative, menschenzentrierte KI-Entwicklung ergänzt, bleiben jedoch in ihrer normativen Ausgestaltung überwiegend abstrakt (Hirsch-Kreinsen & Krokowski, 2024b; Köstler & Ossewaarde, 2022; Ulnicane et al., 2021, 2022). Dies verweist drittens auf ein wiederkehrendes Spannungsverhältnis zwischen deterministischer Innovationsförderung und konstruktivistischer Technikgestaltung und -regulierung, welches bereits in früheren Debatten technologischer Transformation dargelegt wurde (Hirsch-Kreinsen & Krokowski, 2024b; Justo-Hanani, 2022; Mai, 2011; Mäntymäki et al., 2022; Taeihagh, 2021; Weyer, 2008; Willems, 2021).

Schließlich wird viertens eine deutliche Diskrepanz zwischen den bisher vorliegenden Erkenntnissen der politischen Deutung von KI und der techniksoziologischen KI-Kritik sichtbar. Letztere problematisiert strukturelle Aspekte, die scheinbar in der politischen Debatte nur unzureichend adressiert werden. Dazu zählen technokratische Steuerungslogiken, die Verschleierung der sozialen Konstruiertheit von KI, die Reproduktion bestehender Machtungleichgewichte sowie eine solutionistische Reduktion gesellschaftlicher Komplexität (Bartl, 2023; Henneguelle, 2023; Lazar, 2024; Ma, 2020; Sætra, 2020; Weinberg, 2022). Aus dieser Spannung ergeben sich grundlegende Zielkonflikte in Bezug auf Entscheidungsmacht, demokratische Verantwortung und die Legitimität technologiepolitischer Steuerung. Ob und in welcher Form

diese Konflikte im politischen KI-Diskurs aufgegriffen und bearbeitet werden, bleibt bislang jedoch weitgehend offen. Gerade diese Befunde unterstreichen, dass technologiepolitische Steuerungsfragen nur dann legitim und kohärent bearbeitet werden können, wenn die ihnen zugrunde liegenden moralischen Deutungsrahmen explizit gemacht und öffentlich zur Diskussion gestellt werden.

Auf die Bedeutung der moralischen Reflexion verweisen darüber hinaus sowohl die aufgezeigte Ambivalenz und politische Qualität von KI (siehe Abschnitt 6.3.1) als auch ihre drohende *Ex-post-Politisierung* infolge technischer Fehlschläge oder öffentlicher Kontroversen (siehe Abschnitt 6.2.2). Wie dargelegt wurde, wird KI vielfach erst dann problematisiert, wenn gesellschaftliche Schäden bereits eingetreten sind und somit grundlegende Fragen der Legitimität des Technologieeinsatzes sowie nach den zugrunde liegenden normativen Wertvorstellungen aufgeworfen werden (König & Wenzelburger, 2021; Ulbricht, 2020). Eine solche reaktive Politisierung lässt ein strukturelles Defizit prospektiver Legitimitätsproduktion vermuten (Mensch & Schmidt, 2003). Gallego und Kurer (2022) prognostizieren daher eine zunehmende Politisierung von KI-Anwendungen, insbesondere angesichts der weitreichenden Folgen des KI-Einsatzes sowie ungleich verteilter gesellschaftlicher Kosten und Nutzen. Bleibt die moralische Rechtfertigung politischer Entscheidungen jedoch diffus oder technokratisch verkürzt, droht nicht nur ein Legitimationsverlust der Technologie selbst sowie der Politik, sondern auch eine Vertiefung gesellschaftlicher Spaltung zwischen technologischer Elite, skeptischer Öffentlichkeit und marginalisierten Betroffenengruppen.

7.1. Der Bedarf der Rekonstruktion der moralischen Grundlagen politischer Rechtfertigungen Künstlicher Intelligenz

Aus demokratietheoretischer Perspektive erscheint die Untersuchung der moralischen Grundlagen der politischen Rechtfertigungen der KI im gegenwärtigen Entwicklungsstadium daher umso dringlicher. Ausgehend vom Konzept des *Technological Momentum* (Hughes, 1994) befindet sich KI aktuell in einer Phase, in der sich technologische Stabilisierung und gesellschaftliche Dynamik überlagern. Während frühere Phasen der KI-Entwicklung von begrenzten Anwendungsfeldern, technischer Unsicherheit und geringer öffentlicher Aufmerksamkeit geprägt waren, hat sich die Technologie inzwischen zu einem alltagsprägenden und politisch relevanten, wenn auch gering politisierten Phänomen entwickelt (Haenlein & Kaplan, 2019). Durch den breiten Einsatz von KI-Systemen in Bereichen wie Verwaltung, Bildung und Wirtschaft entfalten sich ihre gesellschaftlichen und normativen Implikationen nun mit unmittelbarer Wirkung. In dieser Phase zunehmender institutioneller Verankerung und funktionaler Ausweitung zeigt sich, dass die Technologie nicht nur gestaltet, sondern selbst gestaltend wirkt: Sie prägt

Entscheidungsprozesse, verschiebt Verantwortlichkeiten und beeinflusst grundlegende soziale Ordnungen.

Damit steigt der politische und moralische Rechtfertigungsdruck erheblich. Dieser ergibt sich einerseits aus der sozialen Gestaltbarkeit und Kontingenz technischer Entwicklung, die, einem pragmatischen Technikverständnis folgend, erst in sozialen, politischen und kulturellen Aushandlungen Gestalt annimmt. Politische Entscheidungen über den Einsatz von KI sind somit nie rein technisch, sondern stets normativ fundiert. Angesichts der Unsicherheit über künftige Auswirkungen und der weitreichenden Folgen für zentrale gesellschaftliche Bereiche unterliegen politische Entscheidungsträger:innen einem hohen Rechtfertigungsdruck. Andererseits verstärkt sich dieser Druck durch die normative Wirkmacht technischer Systeme selbst, die Handlungsoptionen strukturieren, soziale Ordnungen stabilisieren oder infrage stellen und damit selbst Teil gesellschaftlicher Machtverhältnisse werden. Technologiepolitik ist in diesem Verständnis grundsätzlich legitimationsbedürftig, da politische Entscheidungen über KI nicht nur mit der Allokation öffentlicher Mittel einhergehen, sondern auch normative Fragen, beispielsweise der persönlichen Autonomie, gesellschaftlichen Verantwortung und sozialen Teilhabe, berühren.

Gerade in dieser kritischen Phase des Übergangs, in der KI weder völlig formbar noch vollständig stabilisiert ist, wird die Frage zentral, auf welchen moralischen Grundlagen ihre Legitimation beruht (Schiff et al., 2022; Tallberg et al., 2023; Züger & Asghari, 2023). Zwar existieren, wie zuvor dargelegt, verschiedene techniksoziologische und technologiepolitische Studien zur Wahrnehmung, Regulierung und strategischen Rahmung von KI, doch eine Analyse der expliziten moralischen Rechtfertigungen in einer Phase beschleunigter gesellschaftlicher Integration steht bislang aus. Wenn historische Technologiekontroversen, etwa um Atomkraft oder Gentechnik, Hinweise liefern, dann zeigt sich, dass in solchen Phasen des *Technological Momentum* die normativen Deutungen darüber, „was Technik darf“ und „wem sie dient“, entscheidend für ihre langfristige Legitimität werden.

Vor diesem Hintergrund ist eine transparente und normativ reflektierte Rechtfertigung politischer Entscheidungen unabdingbar (Brown, 2003; Gallego & Kurer, 2022; Hildebrandt, 2016; Mensch & Schmidt, 2003). Boyd und Holton (2018) betonen die Notwendigkeit, …

> to critically evaluate the range of legitimizing discourses about technological change – whether regarded as inevitable and beneficial, or dangerous and dystopic. This is because stories about technological change do not simply sit above technological innovation and diffusion but help to constitute and direct or re-direct change itself. This widens the

> familiar political-economic emphasis on the sovereign power of the owners of technology to embrace notions of the discursive power of technological rhetoric about social change. (S. 341)

Politische Narrative über technologische Veränderung wirken somit nicht nur legitimierend, sondern tragen aktiv dazu bei, technologische Entwicklungen zu lenken und gesellschaftliche Veränderungen vorzustrukturieren. Politische Rechtfertigungen sind demnach nicht bloße Begleiterscheinungen technischer Innovation, sondern konstituieren deren gesellschaftliche Wirkung mit, indem sie Handlungsoptionen und soziale Ordnungen strukturieren (Perry & Uuk, 2019; Piano, 2020; Taeihagh, 2021). Aus diesem Grund ist es notwendig, die moralischen Grundlagen dieser Rechtfertigungen offenzulegen. Dies gilt nicht nur, um die Reichweite und Grenzen politischer Argumentationen zu verstehen, sondern ebenso, um eine demokratische Kontrolle technologiepolitischer Entscheidungen zu ermöglichen. Auf diese Weise können Legitimationsbedarfe identifiziert sowie mögliche Konflikte prospektiv antizipiert und auf ihre moralisch-normative Qualität hin befragt werden. Vor diesem Hintergrund ergibt sich die zentrale Forschungsfrage (FF) dieser Arbeit:

FF1. Auf welcher moralischen Grundlage wird der Einsatz künstlicher Intelligenz in der politischen Debatte gerechtfertigt?

Zur Beantwortung dieser Frage bietet die SdK einen geeigneten theoretischen Rahmen. Legitimität wird hier als das Ergebnis öffentlicher Prüfungs- und Rechtfertigungsprozesse verstanden, die sich auf diskursiv etablierte Cités (siehe Kapitel 5) im Sinne moralischer Prinzipien stützen (Boltanski & Thévenot, 1991/2018). Diese Cités sind somit geeignet, die Art und Weise, wie politische Akteur:innen ihre Entscheidungen legitimieren, zu analysieren: etwa durch Bezug auf Wettbewerbsvorteile (marktorientierte Cité), auf Effizienzversprechen (industrielle Cité), auf Teilhabe und Gerechtigkeit (staatsbürgerliche Cité) oder auf flexible, netzwerkbasierte Kooperationsformen (projektbasierte Cité).

Die Rekonstruktion der techniksoziologischen Deutungsmuster und empirischen Erkenntnisse der politischen KI-Debatte in Deutschland lässt vermuten, dass markt- und industrieorientierte Rechtfertigungen dominieren, die KI als strategisches Instrument zur Sicherung von Wettbewerbsfähigkeit und Wachstum positionieren, ergänzt durch solutionistische Problemlösungsvisionen. Zugleich kann insbesondere im Lichte partizipativer Technikgestaltung angenommen werden, dass auch staatsbürgerlich orientierte Rechtfertigungen im Sinne sozialer Teilhabe eine Rolle spielen. In welcher Form diese Bezüge tatsächlich mobilisiert werden und wie sie sich zu den dominanten wirtschaftlich-technokratischen Narrativen verhalten, bleibt bislang unklar.

Eine systematische Analyse der zugrundeliegenden Cités ist daher erforderlich, um die moralischen Logiken politischer Rechtfertigung sichtbar zu machen. Diese Untersuchung ist nicht nur für ein vertieftes Verständnis des politischen KI-Diskurses von Bedeutung, sondern auch demokratietheoretisch geboten: Die Legitimität technologiepolitischer Entscheidungen bemisst sich nicht allein an ihrer Effizienz oder Innovationskraft, sondern an der Fähigkeit, unterschiedliche normative Perspektiven zu integrieren, Konflikte offenzulegen und Entscheidungen öffentlich nachvollziehbar zu rechtfertigen. Die Erforschung moralischer Rechtfertigungsprozesse bildet damit eine zentrale Voraussetzung, um die aktuelle Phase der KI-Entwicklung als entscheidenden Moment gesellschaftlicher Sinn- und Wertzuschreibung zu begreifen und die Grundlagen legitimer KI-Politik zu bestimmen.

7.2. Die Analyse der Rolle solutionistischer Rechtfertigungen für die Legitimation Künstlicher Intelligenz

Vor diesem Hintergrund rückt die Analyse solutionistisch-deterministischer Rechtfertigungen in den Blick, die Aufschluss darüber geben kann, welche normativen Verkürzungen und Machtverschiebungen den politischen Umgang mit KI prägen. Deterministische Deutungen bergen das Risiko, technologische Entwicklungen als naturwüchsige und alternativlose Prozesse erscheinen zu lassen. Sie setzen damit die Tradition technokratischer Legitimationsstrategien fort, die moderne Technik als inhärent legitim darstellen, solange ihre Fortschrittsversprechen unwidersprochen bleiben (Habermas, 1968; Schelsky, 1961; Schlaile et al., 2017). Historische Erfahrungen und rekursive Politisierungszyklen zeigen jedoch zum einen, dass solche Rechtfertigungen im Falle technischer Fehlschläge wenig Bestand haben (siehe Abschnitt 6.2.2). Zum anderen verengen sie gesellschaftliche und politische Gestaltungsspielräume und unterschätzen die normativen und sozialen Dimensionen technischer Entwicklung (Wehling, 2021).

Eine derartige fatalistisch-deterministische Haltung droht so, das *Technological Momentum* (Hughes, 1994), also jenen kritischen Moment, in dem Technik noch gestaltbar ist, bevor sie sich zunehmend in gesellschaftlichen Strukturen verfestigt, zu verpassen. Diese technokratische Verengung des politischen Diskurses birgt erhebliche demokratietheoretische Risiken. Wird technologische Entwicklung als sachlogisch notwendig dargestellt, verliert sie ihren politischen Aushandlungscharakter. Entscheidungen erscheinen dann nicht mehr als kontingente Wahl zwischen Alternativen, sondern als unvermeidliche Folge „objektiver“ technologischer Notwendigkeiten. Dieses Narrativ unterläuft demokratische Legitimität, indem politische Gestaltung auf die Verwaltung vermeintlicher Sachzwänge reduziert und pluralistische Partizipation marginalisiert werden (Bartl, 2023; Hay, 2007; Köstler & Ossewaarde, 2022).

Die Verbreitung solutionistischer Problemlösungsversprechen ist dabei eingebettet in eine zunehmend netzwerkbasierte Innovationspolitik, die stark von privatwirtschaftlichen und akademischen Akteurskonstellationen geprägt ist (Blümel, 2021; Bogner, 2021b; Heyen et al., 2024; Mai, 2011). Dadurch wird politische Verantwortung diffuser, parlamentarische Kontrolle geschwächt und normative Aushandlungsprozesse werden zunehmend von ökonomischen Rechtfertigungslogiken überlagert (Khanal et al., 2025; Lindman et al., 2023). Dies verschärft Machtungleichgewichte zwischen Staat, Wirtschaft und Zivilgesellschaft und erschwert demokratische Rechenschaftspflicht. Besonders problematisch ist dies, wenn politische Entscheidungen kurzfristig auf technologische Impulse reagieren, ohne langfristige soziale Folgen zu reflektieren. So drohen technologische Gestaltungsperspektiven verkürzt, Entscheidungsprozesse entdemokratisiert und Mitbestimmungsmöglichkeiten eingeschränkt zu werden (Bartl, 2023; Henneguelle, 2023; Ma, 2020; Salais, 2016).

Umso wichtiger ist es, den politischen Umgang mit KI nicht auf technische Machbarkeit oder ökonomische Verwertbarkeit zu reduzieren, sondern ihn im Lichte techniksoziologischer und demokratietheoretischer Erkenntnisse kritisch zu reflektieren. Eine Analyse solutionistischer Rechtfertigungen kann dazu beitragen, implizite Deutungsmuster, Machtasymmetrien und normative Leerstellen aufzudecken. Denn solutionistische Denkmuster tragen nicht nur zur technokratischen Verengung politischer Entscheidungsprozesse bei, sondern fördern zugleich die Reproduktion elitärer Deutungshoheit. Vor dem Hintergrund dieser demokratietheoretisch relevanten Gefahren verdient der Solutionismus besondere analytische Aufmerksamkeit. Die Analyse richtet sich daher auf die Aufdeckung solutionistischer Rechtfertigungen im politischen Diskurs über KI und greift damit ein zentrales Forschungsdesiderat auf (Nachtwey & Seidl, 2023):

> *FF2. Inwiefern manifestieren sich solutionistische Rechtfertigungen in der politischen Debatte über die Einführung Künstlicher Intelligenz?*

Obwohl sich in der Wissenschaft eine Distanzierung vom Technikdeterminismus durchgesetzt hat, deuten empirische Befunde darauf hin, dass Technik im öffentlichen Diskurs weiterhin als exogener Innovationsfaktor gedeutet wird, der möglichst effizient in gesellschaftliche Prozesse integriert werden soll. Deterministische Deutungsmuster strukturieren mitunter die politische KI-Debatte, die als elitengetriebenes Projekt aufgefasst wird, das zentrale gesellschaftliche Gruppen faktisch von der Mitgestaltung ausschließt: „AI remains an elitist project steered by political-administrative and corporate interests“ (Köstler & Ossewaarde, 2022, S. 261). Zum Ausdruck kommen diese Deutungsmuster in der Rhetorik der Alternativlosigkeit, im

verstärkten Rückgriff auf Expert:innengremien oder in solutionistischen Versprechen, gesellschaftliche Herausforderungen wie Bildungsgerechtigkeit, Gesundheit oder Bürokratieabbau durch KI effizienter zu bewältigen (Hjaltalin & Sigurdarson, 2024; Mai, 2011; Morozov, 2013; Ossewaarde & Gülenç, 2020). Dieses Narrativ wird insbesondere durch die sogenannte Digital-Elite befeuert, die die gesellschaftlichen Problemlösungspotenziale der KI betont (Brockmann et al., 2021; Nachtwey & Seidl, 2023; Obendiek & Seidl, 2023; Stilgoe, 2020).

Angesichts des hohen Vertrauens politischer Akteur:innen in Expert:innen, insbesondere in Zeiten der Unsicherheit, sowie des starken Einflusses großer Technologiekonzerne auf Technikentwicklung und auch -regulierung (Cath, 2018; Khanal et al., 2025; Lindman et al., 2023; Sharon, 2021) kann daher angenommen werden, dass diese die wirtschaftlich und solutionistisch geprägten Rechtfertigungen der Tech-Elite übernehmen könnten. Morozov (2020) geht gar davon aus, dass die Ideologien des Neoliberalismus und des Solutionismus staatliche Institutionen bereits jetzt so tiefgreifend durchdrungen haben, dass diese zu *den* handlungsleitenden Narrativen der Gegenwart werden. Zugleich wurde im vorausgehenden Kapitel jedoch bereits die Ambivalenz der Steuerung technologiepolitischer Prozesse dargelegt: Einerseits betont die Bundesregierung mit dem Leitbild einer partizipativen Technikgestaltung die Einbindung verschiedener Stakeholder, um eine aktive und gemeinwohlorientierte KI-Entwicklung zu fördern. Andererseits zeigen kritische Analysen, dass diese Partizipationsversprechen häufig formaler Natur bleiben. Insbesondere die zunehmende Institutionalisierung von Expert:innennetzwerken, etwa durch die Datenethikkommission, geht mit einem Bedeutungsverlust parlamentarischer Debatten einher (Willems, 2021). Aufgrund dieser Uneindeutigkeit der bislang vorliegenden empirischen Erkenntnisse rückt die Frage nach der demokratischen Verfasstheit technologiepolitischer Steuerung erneut ins Zentrum. Dabei wird die Rolle solutionistischer Rechtfertigungen in der politischen Legitimation von KI kritisch analysiert.

7.3. Die Untersuchung der kritischen Fundierung der Legitimation Künstlicher Intelligenz

Neben der Bedenklichkeit solutionistischer Rechtfertigungen des KI-Einsatzes liegt eine weitere Gefahr in einer Depolitisierung der politischen KI-Debatte, die durch weitgehende Konfliktfreiheit droht, in einer tautologischen Selbstbestätigung politischer Akteur:innen zu münden (Boltanski, 2010). So verdeutlichen die Erfahrungen früherer Technologiepolitik, dass technologische Hypes kritisch zu hinterfragen sind. Stattdessen erfordert die politische Legitimität technischer Entwicklung nicht ausschließlich einen Modus der Zustimmung oder Stabilität, sondern sie konstituiert sich erst im Modus der Kritik und des Konflikts, wie etwa kritisch-pragmatische Technikperspektiven ebenso betonen wie die SdK. Erst durch die Möglichkeit

der Infragestellung und des Widerstands wird politische Legitimation überprüfbar und verhandelbar. In der Begrifflichkeit der SdK markieren insbesondere die Realitäts- und Existenzprüfung zentrale Momente legitimatorischer Aushandlung (Boltanski, 2010): Politische Rechtfertigungen müssen sich daran messen lassen, ob sie als glaubwürdig, umsetzbar und gesellschaftlich tragfähig anerkannt werden oder ob sie auf Skepsis, Ablehnung oder fundamentale Gegenargumente stoßen. Ohne die Möglichkeit solcher Prüfungen verkommt politische Rechtfertigung zu bloßer Legitimationsrhetorik. Kritik und Konflikt sind nicht nur Begleiterscheinungen politischer Entscheidungsprozesse, sondern konstitutive Voraussetzungen demokratischer Legitimität. Sie ermöglichen die Artikulation abweichender Interessen, eröffnen Aushandlungsräume und fördern die Reflexivität politischer Entscheidungen. Nicht zuletzt bilden sie das Fundament für eine politische Öffentlichkeit, die Entscheidungen hinterfragen, Alternativen formulieren und Rechenschaft einfordern kann. Kritik gilt somit als jenes Korrektiv, das eine demokratische Gestaltung von Technik erst ermöglicht.

Vor diesem Hintergrund erscheint es umso bedeutsamer, dass die bislang vorliegenden empirischen Erkenntnisse auf eine augenscheinliche Abwesenheit solcher konfliktiven Aushandlungen in der politischen KI-Debatte in Deutschland hinweisen. Der Diskurs scheint bislang weitgehend konfliktarm und positiv konnotiert zu verlaufen sowie primär auf die Darstellung von Potenzialen abzuzielen. Hinweise auf das Ausloten normativer Spannungsfelder, wie etwa im Bereich des Datenschutzes, tauchen bislang nur vereinzelt auf (Guenduez & Mettler, 2023; Hirsch-Kreinsen & Krokowski, 2024b; Lambach & Oppermann, 2023; Willems, 2021). In der politischen KI-Debatte scheint Legitimität somit häufig implizit aus dem erwarteten Nutzen von KI abgeleitet zu werden, der zumeist wirtschaftlich definiert ist und dessen Nebenwirkungen unbekannt sind.

Kritischere Perspektiven scheinen hingegen bislang weitestgehend in der akademischen, kritischen techniksoziologischen Debatte zu verbleiben (Eynon & Young, 2021). In der politischen Debatte sind diese hingegen meist unterrepräsentiert, erscheinen vereinzelt oder finden nur dann Eingang, wenn unmittelbar politisch anschlussfähige Lösungsvorschläge unterbreitet werden können, wie zum Beispiel in Form ethischer Prinzipien oder technischer Standards (Köstler & Ossewaarde, 2022; Ulnicane et al., 2021). Dieser Modus der Kritikverarbeitung läuft Gefahr, weniger zur Aushandlung kontroverser Positionen als vielmehr zur symbolischen Entschärfung normativer Einwände beizutragen. Die Thematisierung von Risiken wie Diskriminierung, Datenschutz oder algorithmischer Intransparenz erfolgt somit häufig in Form technischer Kontrollierbarkeit, wobei ihre gesellschaftliche Dimension hingegen oft unadressiert bleibt. Ebenso verweist die *Ex-post-Politisierung* infolge konkreter Schäden (König & Wenzelburger, 2021;

Ulbricht, 2020) auf ein grundsätzlich reaktives Politikverständnis, das nicht auf prospektive Normenbildung, sondern auf Schadensbegrenzung ausgerichtet ist und so die Möglichkeit einer demokratisch-partizipativen Technikgestaltung einschränkt. Einem solchen Verständnis mangelt es an einer institutionalisierten Konfliktkultur, die kritische Perspektiven nicht als Abweichung, sondern als notwendige Bestandteile legitimer Entscheidungsfindung begreift.

Vor diesem Hintergrund liegt auf Basis der bisherigen Erkenntnisse die Vermutung nahe, dass die gegenwärtige politische Auseinandersetzung mit KI nur sehr eingeschränkt die Bedingungen legitimer Entscheidungsfindung erfüllt. Explizit oder implizit auftretende moralische Konfliktlinien sind bislang jedoch kaum systematisch untersucht worden. Damit bleiben zentrale gesellschaftliche Spannungsverhältnisse, etwa zwischen technischer Innovation und sozialer Gerechtigkeit oder zwischen demokratischer Kontrolle und technokratischer Steuerung, weitgehend unterbeleuchtet. Darüber hinaus liefe eine auf die bloße Rechtfertigung beschränkte Analyse Gefahr, die aus demokratietheoretischer Perspektive notwendige Rolle der Kritik und des Konflikts zu übersehen. Diese Leerstelle markiert somit ein zentrales Desiderat für die Forschung und begründet die abschließende Forschungsfrage:

> *FF3. Auf welcher moralischen Grundlage wird der Einsatz von KI in der politischen Debatte kritisiert und welche Konfliktlinien treten dabei zutage?*

Die Beantwortung dieser Frage ermöglicht, die Konflikt- und Kritikfähigkeit des politischen Systems im Umgang mit KI zu beurteilen. Erst das Zusammenspiel von Kritik und Rechtfertigung erlaubt Rückschlüsse auf das Ausmaß und die Qualität der Politisierung technischer Innovationen sowie auf die Frage, ob und wie demokratische Legitimation durch kritische Infragestellung, Aushandlungsprozesse und normative Prüfung ermöglicht oder verhindert wird. Insofern bildet diese Untersuchung einen wesentlichen Bestandteil der Gesamtanalyse der politischen KI-Debatte. Dies gilt nicht trotz, sondern gerade aufgrund der augenscheinlichen Konfliktlosigkeit.

8. Methodisches Vorgehen

Im Zentrum der folgenden Untersuchung stehen somit die Rechtfertigung und die Kritik des KI-Einsatzes im deutschen Bildungssektor, die als zentrale Grundlage politischer Legitimationsbemühungen verstanden werden. Dabei stützt sich die Arbeit auf eine Dokumentenanalyse politischer Kommunikation, die eine systematische Dokumentenauswahl mit einer inhaltlich strukturierenden qualitativen Inhaltsanalyse verbindet. Diese Kombination ermöglicht eine hohe Generalisierbarkeit der Ergebnisse und macht zugleich neben manifesten auch latente Kommunikationsinhalte sichtbar (Kuckartz & Rädiker, 2022). Auf diese Weise lassen sich die

Rechtfertigungsprozesse des KI-Einsatzes im Bildungssystem systematisch erfassen und die Aushandlung von Legitimität nachzeichnen. Dabei sichert die Triangulation unterschiedlicher Methoden und Datenquellen eine hohe externe Studiengüte (Denzin, 1970/2017; Lamnek & Krell, 2016). Die ausführliche nachfolgende Darstellung des methodischen Vorgehens dient darüber hinaus der intersubjektiven Nachvollziehbarkeit und ermöglicht die Bewertung der internen Studiengüte (Creswell & Creswell Báez, 2021; Flick, 2020; Kirk & Miller, 1989; Kuckartz & Rädiker, 2022; Mayring, 2016). Im Folgenden werden der Untersuchungsgegenstand definiert, die Methodik der Dokumentenrecherche und -auswahl sowie der qualitativen Inhaltsanalyse erläutert und die weiteren Mechanismen zur Sicherung der qualitativen Studiengüte dargestellt.

8.1. Definition des Untersuchungsgegenstands

Zur Definition des Untersuchungsgegenstandes gilt es zunächst zu klären, an welchem Ort die Auseinandersetzung über den Einsatz von KI im Bildungswesen geführt wird und wo in der Konfrontation unterschiedlicher Positionen Legitimität erzeugt werden soll. Dabei manifestiert sich die Grundlage politischer Entscheidungsprozesse überwiegend in schriftlicher Form, etwa in Policy-Dokumenten, protokollierten Parlamentsdebatten, Richtlinien oder Stellungnahmen. Es liegt daher nahe, diese Dokumente in den Mittelpunkt der Analyse zu stellen. In ihnen wird sichtbar, welche Rechtfertigungen von politischen Akteur:innen sowie von zivilgesellschaftlichen und wissenschaftlichen Beteiligten in politische Prozesse eingebracht werden, wie Positionen ausgehandelt werden, welche Konfliktlinien entstehen und schließlich wie Legitimität für oder gegen den Einsatz von KI erzeugt werden soll. Ihre Analyse ermöglicht es somit, die in politischen Stellungnahmen erkennbaren Ziele und Werte politischer Entscheidungen zu rekonstruieren (Cardno, 2018) und sie mit den zugrunde liegenden Cités in Verbindung zu bringen. Zugleich ist zu berücksichtigen, dass die Legitimation politischer Entscheidungen nicht allein durch den Aushandlungsprozess bestimmt wird, sondern ebenso durch die gesellschaftliche Wahrnehmung der Angemessenheit politischen Handelns geprägt ist (siehe Abschnitt 3.1). Vor diesem Hintergrund fokussiert die Analyse auf jene Dokumente, die im Unterschied zu informellen oder lobbyistischen Diskursen der breiten Öffentlichkeit prinzipiell zugänglich sind.

Um die politische Debatte möglichst umfassend abzubilden und der Politisierungsdimension einer breiten Reichweite der Akteur:innen gerecht zu werden (siehe Abschnitt 4.3), werden neben parlamentarischen Materialien auch Stellungnahmen jener Institutionen berücksichtigt, die politische Entscheider:innen beraten oder als zivilgesellschaftliche Akteur:innen unmittelbar vom KI-Einsatz im Bildungswesen betroffen sind. Von diesen Gruppen ist anzunehmen, dass sie einen erheblichen Einfluss auf politische Entscheidungen ausüben können. Berücksichtigt

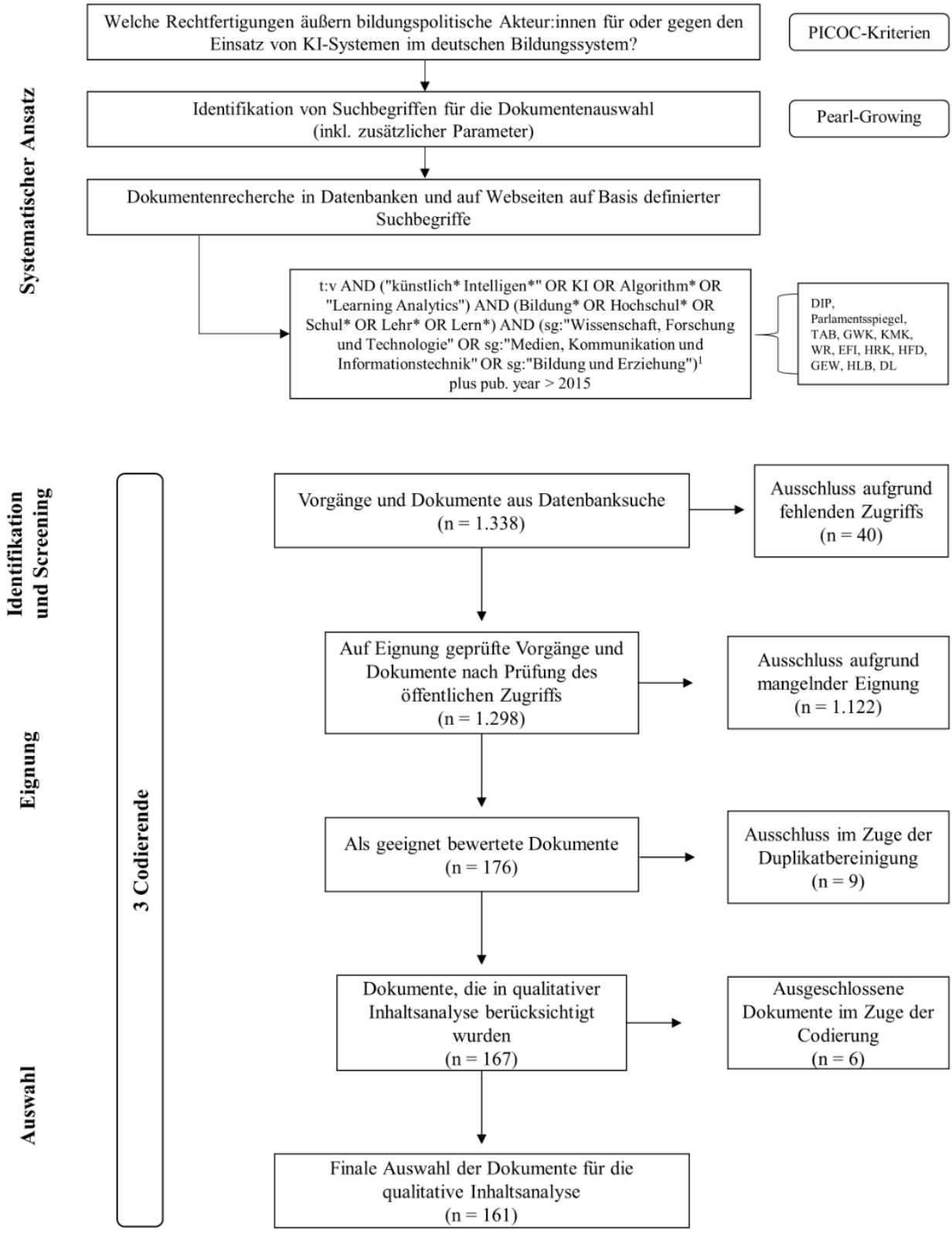
werden daher sowohl Policy-Dokumente und Parlamentsdebatten als auch Publikationen, Pressemitteilungen und Blogbeiträge beratender oder zivilgesellschaftlicher Institutionen. Neben dem Deutschen Bundestag (BT), dem Bundesrat (BR) und den 16 Landtagen (LT) als Legislativinstitutionen werden zum einen mit der Gemeinsamen Wissenschaftskonferenz (GWK) und der Kultusministerkonferenz (KMK) weitere politische, den Bildungsbereich betreffende bundländer-koordinierende Gremien einbezogen. Zum anderen werden mit dem Wissenschaftsrat (WR), der Expertenkommission Forschung und Innovation (EFI) und dem Büro für Technikfolgenabschätzung (TAB) bildungs- beziehungsweise technologiepolitisch relevante Beratungsgremien berücksichtigt. Schließlich werden mit der Hochschulrektorenkonferenz (HRK), dem Hochschullehrerbund (HLB), der Gewerkschaft Erziehung und Wissenschaft (GEW), dem Deutschen Lehrerverband (DL) und dem Hochschulforum Digitalisierung[17] (HFD) bildungsspezifische und gewerkschaftliche Institutionen in die Untersuchung aufgenommen. Nachfolgend wird die Auswahl der einzubeziehenden Dokumente dieser Institutionen erläutert.

8.2. Dokumentenrecherche und -auswahl

Die Methodenliteratur der empirischen Sozialforschung bietet verschiedene Ansätze für die inhaltsanalytische Untersuchung von Dokumenten (Salheiser, 2022). Dabei verweisen einige Autor:innen auf geeignete Datenbanken zur Recherche (z. B. Hoffmann, 2018) und beschreiben unterschiedliche Formen der Stichprobenziehung (z. B. Salheiser, 2022), wobei jedoch häufig offen bleibt, wie die Auswahl des tatsächlichen Untersuchungsgegenstands systematisch begründet wird. Die Dokumentenanalyse verfügt damit über keine klar etablierte Vorgehensweise zur Auswahl der einzubeziehenden Dokumente, obwohl dieser Schritt für die wissenschaftliche Qualität der Ergebnisse zentral ist. Um einer willkürlich wirkenden Auswahl entgegenzuwirken und eine nachvollziehbare Begründung des Einbezugs der Dokumente zu ermöglichen, wurden daher verschiedene Verfahren der Literaturauswahl systematischer Reviews nach Booth et al. (2016), Higgins et al. (2019) sowie Petticrew und Roberts (2006) kombiniert und adaptiert. Auf dieser Grundlage ergibt sich ein vierstufiges Vorgehen: Erstens die Definition der Forschungsfrage, zweitens die Festlegung von Einschlusskriterien, drittens eine umfassende Recherche und viertens die Sichtung und eigentliche Auswahl der Ergebnisse (Booth et al., 2016; Higgins et al., 2019; Petticrew & Roberts, 2006). Die Adaption dieser Schritte bildet die Grundlage des im Folgenden dargestellten Vorgehens. Die Dokumentenrecherche und -auswahl werden in einem an Moher et al. (2009) angelehnten PRISMA-Flowchart in Abbildung 2 visualisiert.

[17] Das HFD ist ein Think-&-Do-Tank auf Initiative des Stifterverbandes, des Centrums für Hochschulentwicklung und der HRK zur Beratung und Vernetzung von Hochschulen, Politik, Wirtschaft & Gesellschaft, welches durch das Bundesministerium für Forschung, Technologie und Raumfahrt seit 2014 gefördert wird (HFD, 2025).

Abbildung 2. Flowchart zur Auswahl der Untersuchungseinheiten in Anlehnung an Moher et al. (2009)



[1] Hier: DIP Suchstring; Suchstring für Parlamentsspiegel.de geringfügig an Datenbank-spezifische Anforderungen angepasst; andere Institutionen über Webseiten-Suche basierend auf erstem Bestandteil des Suchstrings aufgrund mangelnder Option für erweiterte Suche

8.2.1. Strukturierung der Forschungsfrage

Die Dokumentenanalyse verfolgt das Ziel, die inhaltlichen Rechtfertigungen des Einsatzes von KI-Systemen zu untersuchen. Die daraus folgenden Forschungsfragen wurden bereits aus dem Erkenntnisinteresse der Arbeit abgeleitet (siehe Kapitel 7). Für die Dokumentenrecherche wurde die zugrunde liegende Frage jedoch im ersten Schritt erneut hinsichtlich ihrer analytischen Präzision und ihrer Eignung zur Ableitung von Suchbegriffen überprüft. Dafür wurde die PICOC-Methode herangezogen, um die zentralen Elemente der Fragestellung systematisch zu bestimmen (Booth et al., 2016; Petticrew & Roberts, 2006). Dabei steht PICOC für die Definition von **P**opulation, **I**ntervention beziehungsweise Exposure, **C**omparison, **O**utcome und **C**ontext und ermöglicht folgende Bestimmungen (siehe Tabelle 3): Erstens bilden bildungspolitische Entscheider:innen und Berater:innen die Population. Zweitens bezieht sich die Intervention auf den Einsatz von KI-Systemen im Bildungssektor. Drittens wird keine vergleichende Analyse angestrebt, weshalb dieses Element entfällt. Viertens definieren die geäußerten Rechtfertigungen für oder gegen den Einsatz von KI das Outcome der Analyse. Fünftens wird der Kontext auf das deutsche Bildungssystem respektive die bildungspolitische Debatte eingegrenzt. Daraus ergibt sich folgende konkretisierte Forschungsfrage für die Dokumentenanalyse: Welche Rechtfertigungen äußern bildungspolitische Akteur:innen[18] für oder gegen den Einsatz von KI-Systemen im deutschen Bildungssystem?

Tabelle 3: Anwendung der PICOC-Methode nach Petticrew & Roberts (2006)

P	–	opulation	Bildungspolitische Berater:innen & Entscheider:innen
I	–	ntervention	Einführung von KI-Systemen
C	–	omparison	–
O	–	utcome(s)	Rechtfertigungen (für oder gegen den Einsatz)
C	–	ontext	Deutsches Bildungssystem

8.2.2. Definition von Einschlusskriterien: Untersuchungszeitraum, Datenbanken und Suchbegriffe

Darauf aufbauend wurden im zweiten Schritt Einschlusskriterien für die Dokumentenauswahl definiert. Dazu gehören die Festlegung des Untersuchungszeitraums, der einzubeziehenden Datenbanken und der Suchbegriffe. Für die Analyse wurden alle öffentlich verfügbaren Dokumente der ausgewählten bildungspolitischen Institutionen berücksichtigt, die im Zeitraum vom 1. Januar 2015 bis zum 31. Dezember 2022 veröffentlicht wurden. Zwar beginnt die Entwicklung von „Artificial Intelligence in Education“ und Learning Analytics bereits in den frühen

[18] An dieser Stelle wird auf die Präzisierung der Akteur:innenbeschreibung als bildungspolitische Berater:innen und Entscheider:innen verzichtet, da diese für die Festlegung der nachfolgend vorgestellten Suchkriterien nicht von Bedeutung sind und bereits durch die Definition des Untersuchungsgegenstandes abgedeckt wird.

2010er Jahren (Ferguson et al., 2016), sie konzentriert sich jedoch zunächst auf den englischsprachigen Raum. Im deutschen Bildungssektor setzte die wissenschaftliche und politische Auseinandersetzung später ein (Ifenthaler & Schumacher, 2016a; Keller et al., 2019), was sich unter anderem an der ersten Richtlinie zur Förderung von Forschung zur digitalen Hochschulbildung im Jahr 2016 (BMBF, 2016) und an einer Diskursanalyse des KI-Begriffs im Deutschen Bundestag zeigt (Willems, 2021).

Hinsichtlich der verwendeten Datenbanken wurde für Bundestag und Bundesrat das gemeinsame Dokumentations- und Informationssystem für Parlamentsmaterialien (DIP) genutzt. Plenarprotokolle und Drucksachen der Landtage wurden über das Informationssystem Parlamentsspiegel.de recherchiert. In allen anderen Fällen bildeten die Webseiten der jeweiligen Institutionen die Grundlage der Dokumentenrecherche. Hier wurde auf institutionseigene Datenbanken, das Archiv oder die einfache Suchfunktion der Webseite zurückgegriffen.[19]

Für die Ableitung konkreter Suchbegriffe bildet die Forschungsfrage den Ausgangspunkt. So wurden zunächst die Begriffe „Künstliche Intelligenz", „Bildung" und „Einführung" als Suchkriterien festgelegt. Der Begriff „Rechtfertigung" wurde nicht aufgenommen, da Rechtfertigungen zwar das theoretische Analysekonstrukt darstellen, jedoch nicht zwingend explizit als solche benannt werden. Die spätere qualitative Inhaltsanalyse soll daher gerade auch implizit geäußerte Rechtfertigungen erfassen. Zur Präzisierung und Erweiterung der Suchbegriffe wurde anschließend die Pearl-Growing-Methode angewendet. Sie dient dazu, auf Basis eines im Vorfeld identifizierten Dokuments, welches einen wesentlichen Beitrag zur Beantwortung der Forschungsfrage liefert, weitere Suchbegriffe abzuleiten (Booth et al., 2016). Mithilfe der zunächst definierten Suchbegriffe wurde in der DIP-Datenbank der Antrag „Smart Germany – Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen" der FDP-Bundestagsfraktion vom 16.10.2019 identifiziert. Auf dieser Grundlage wurden zusätzliche Suchbegriffe ergänzt, die entweder synonym zum KI-Begriff verwendet werden (Komponente I) oder den Bildungskontext weiter präzisieren (Komponente II). Dabei wurden die Elemente der Intervention und des Kontexts (**PICO**) als strukturierende Elemente zweier Komponenten herangezogen.[20] Ergänzt wurden die Begriffe „KI", „Algorithmus" und „Learning Analytics" sowie „Hochschule", „Schule", „Lehren" und „Lernen" (siehe Tabelle 4). Gleichzeitig wurde „Einführung" als Suchbegriff gestrichen, da die Pearl-Growing-Analyse zeigte, dass zu voraussetzungsvoll definierte Begriffe relevante Dokumente irrtümlicherweise ausschließen könnten.

[19] Die konkreten Suchstrategien für jede einzelne Institution können Anhang A2 „Codebuch der Dokumentenauswahl" sowie der Übersicht in Tabelle 5 entnommen werden.
[20] Begriffe rund um bildungspolitische Berater:innen und Entscheider:innen wurden nicht als Suchkriterien aufgenommen, da diese bereits über die zu untersuchenden Institutionen festgelegt wurden.

Tabelle 4: Suchbegriffe der Dokumentenrecherche

Komponente I	**Komponente II**
„Künstlich* Intelligen*“	Bildung*
KI	Hochschul*
Algorithm*	Schul*
„Learning Analytics“	Lehr*
	Lern*

8.2.3. Dokumentenrecherche

Der dritte Schritt umfasste die Dokumentenrecherche. Dafür wurde zunächst eine Lemmatisierung der festgelegten Suchbegriffe vorgenommen. Die Begriffe wurden mithilfe von Trunkierungen (*) auf ihre Stammform reduziert, um sämtliche Wortvarianten in die Suche einzubeziehen, sodass etwa nicht nur „Algorithmus“ erfasst wird, sondern ebenso „algorithmisch“, „Algorithmen“ oder „algorithmischer“. Anschließend wurden die Suchbegriffe mithilfe boolescher Operatoren zu zwei kombinatorischen Einheiten zusammengeführt und zu einem Suchstring verbunden (Lefebvre et al., 2019). Dieser bildete neben der Einschränkung des Untersuchungszeitraums die Grundlage der erweiterten Suche im Parlamentsspiegel sowie für die Expertensuche im DIP, die an dieser Stelle exemplarisch erläutert wird. So wurde die Recherche im DIP auf Vorgänge beschränkt, wodurch thematisch zusammengehörige Materialien gebündelt werden, wie zum Beispiel Anträge und dazugehörige Parlamentsdebatten. Auf diese Weise konnten sowohl eine forschungsökonomisch bearbeitbare Trefferanzahl gewährleistet (9.813 Einzeldokumente gegenüber 1.758 Vorgängen, Stand 15.08.2022) als auch sichergestellt werden, dass alle Dokumente eines im späteren Auswahlprozess als relevant bewerteten Vorgangs berücksichtigt werden können. Zudem wurde die Suche auf die drei Sachgebiete (1) Wissenschaft, Forschung und Technologie, (2) Medien, Kommunikations- und Informationstechnik sowie (3) Bildung und Erziehung eingegrenzt, um eine thematisch sinnvolle Zuordnung des Hauptthemas des Vorgangs sicherzustellen. Daraus ergibt sich der folgende Suchstring, wobei der Untersuchungszeitraum zusätzlich manuell als Filteroption erfasst wurde:

```
t:v AND („künstlich* Intelligen*“ OR KI OR Algorithm* OR „Learning Analytics“) AND (Bildung* OR Hochschul* OR Schul* OR Lehr* OR Lern*) AND (sg:“Wissenschaft, Forschung und Technologie“ OR sg:“Medien, Kommunikation und Informationstechnik“ OR sg:“Bildung und Erziehung“)
```

Für den Parlamentsspiegel sowie die Webseite des TAB wurden entsprechend angepasste Suchstrings verwendet. Auf den Webseiten der übrigen neun Institutionen wurden ausschließlich die Suchbegriffe der ersten Komponente beziehungsweise im Fall der EFI der zweiten Komponente

herangezogen, da deren thematische Ausrichtung bereits eine Eingrenzung auf den Bildungssektor (respektive einen Innovationsfokus) gewährleistet. Eine vollständige Übersicht der Suchanfragen und Trefferzahlen findet sich in Tabelle 5.

Die Dokumentenrecherche wurde durch drei Forschende[21] auf Basis eines umfänglichen Codebuchs durchgeführt (siehe Anhang A2) und detailliert dokumentiert. Die Erfassung der Suchtreffer erfolgte in einer Excel-Tabelle (siehe Anhang B2), die neben einer fortlaufenden Nummer die formalen Kategorien Institution, Dokumententitel, Abstract, Erscheinungsdatum, Dokumententyp (z. B. Plenarprotokoll, Pressemitteilung, Blogbeitrag) sowie den Zugangslink[22] erfasst. Im Untersuchungszeitraum wurden insgesamt n = 1.338 Vorgänge und Einzeldokumente identifiziert. Die meisten Treffer entfielen auf Parlamentsmaterialien: Im DIP wurden n = 467 Vorgänge des Bundestags und Bundesrats erfasst, in den Landtagen n = 590 Dokumente. Damit stammen 79 % aller Suchtreffer aus parlamentarischen Quellen. Der HLB wies als einzige Institution keine Treffer auf Grundlage der Webseitensuche auf und auch für den DL und die HRK wurden lediglich zwei respektive fünf Dokumente ermittelt, die sich jedoch im nachfolgenden Auswahlschritt als nicht relevant erwiesen.

8.2.4. Dokumentenauswahl

Die abschließende systematische Dokumentenauswahl folgte einem vierstufigen Vorgehen. Dabei beurteilten drei Forschende die Relevanz der Dokumente im Hinblick auf die Beantwortung der Forschungsfrage und identifizierten so jene Dokumente, die als Untersuchungseinheiten dienen. Eine umfassende Schulung auf Grundlage eines Codebuchs (siehe Anhang A2) stellte gemeinsam mit der Prüfung der Intercoderreliabilität vor Beginn des Auswahlprozesses ein konsistentes Vorgehen sicher. In drei Durchläufen wurden von den insgesamt ermittelten n = 1.338 Suchtreffern n = 131 Dokumente zufällig ausgewählt und von allen Codierenden beurteilt. Die Ergebnisse wurden nach jedem Durchlauf diskutiert, um die Übereinstimmung zwischen den Codierenden qualitativ zu verbessern. Zur Bestimmung der Intercoderreliabilität wurde aufgrund der nominalskalierten Daten die prozentuale Übereinstimmung nach Holsti (1969) berechnet. Die Reliabilität verbesserte sich von r_H = .835 im ersten auf r_H = .960 im dritten Durchlauf (siehe Anhang A3), sodass eine sehr gute Intercoderreliabilität für die Auswahl der Dokumente gewährleistet wurde (Lombard et al., 2002).

[21] Neben der Autorin dieser Arbeit wurden die Dokumentenrecherche sowie die qualitative Inhaltsanalyse jeweils durch zwei Hilfskräfte des Lehrstuhls Kommunikations- und Medienwissenschaft I der Heinrich-Heine-Universität Düsseldorf unterstützt.

[22] Die Dokumente wurden zu diesem Zeitpunkt aus forschungsökonomischen Gründen noch nicht heruntergeladen. Ihre Reidentifikation ist daher abhängig von der dauerhaften Verfügbarkeit der dokumentierten Titel und Zugriffslinks (siehe Anhang B2).

Tabelle 5. Suchabfragen, Suchtreffer und Untersuchungseinheiten nach Institutionen

Institution	Datenbank/ Suchfunktion	Publikationstyp	Suchbegriffe	Suchtreffer	Auswahl
Bundestag/ Bundesrat	DIP: Expertensuche	Drucksachen & Plenarprotokolle	t:v AND ("künstlich* Intelligen*" OR KI OR Algorithm* OR "Learning Analytics") AND (Bildung* OR Hochschul* OR Schul* OR Lehr* OR Lern*) AND (sg:"Wissenschaft, Forschung und Technologie" OR sg:"Medien, Kommunikation und Informationstechnik" OR sg:"Bildung und Erziehung")	467 (34,9 %)	28 (17,4 %)
Landtage	Parlamentsspiegel.de: Erweiterte Suche	Drucksachen & Plenarprotokolle	Künstliche Intelligen*' OR 'Künstlich Intelligent*' OR 'Künstlicher Intelligenz' OR KI OR Algorithm* OR 'Learning Analytics' OR 'HOCHSCHULEN' OR 'HOCHSCHULEN FÃ?R ANGEWANDTE WISSENSCHAFTEN' OR 'HOCHSCHULEN FÜR ANGEWANDTE WISSENSCHAFTEN' OR 'HOCHSCHULWESEN' OR 'BILDUNG' OR 'SCHULEN' OR 'TECHNOLOGIE'	590 (44,1 %)	53 (32,9 %)
GWK	Webseite: (I) Publizierte Materialien > Volltextsuche + (II) Webseitensuche	Dokumente der GWK	Künstl* - KI - Algorithm* - Learning Analytics	29 (2,2 %)	15 (9,3 %)
KMK	Webseite: (I) Beschlüsse und Veröffentlichungen > Bildung in der digitalen Welt + (II) Publizierte Materialien > Volltextsuche + Webseitensuche + (III) Pressearchiv	Beschlüsse & Veröffentlichungen	Künstl* - KI - Algorithm* - Learning Analytics	16 (1,2 %)	2 (1,2 %)
WR	Webseite: Publikationssuche	Publikationen	Künstl* - KI - Algorithm* - Learning Analytics	98 (7,3 %)	17 (10,6 %)
EFI	Webseite: (I) Gutachten + (II) Policy Briefs	Gutachten & Policy Briefs	Bildung, Hochschul*, Schul*. Lehr*, Lern*	9 (0,7 %)	2 (1,2 %)
TAB	Website: (I) Publikationssuche + (II) TABlog	Publikationen, Blogbeiträge	Künstlich* - Algorithm* - Learn* Bildung - Schul* - Lern* - Lehr*	14 (1,0 %)	7 (4,4 %)
HRK	Webseite: (I) Publikationssuche + (II) Webseitensuche	Publikationen, Pressemitteilungen & Beschlüsse	Künstl* - KI - Algorithm* - Learning Analytics	5 (0,4 %)	0 (0,0 %)
HLB	Webseite: Webseitensuche	Dokumente	Künstl* - KI - Algorithm* - Learning Analytics	0 (0,0 %)	0 (0,0 %)
GEW	Webseite: Webseitensuche	Publikationen, Pressemitteilungen, Nachrichten	Künstl* - KI - Algorithm* - Learning Analytics	78 (5,8 %)	19 (11,8 %)
DL	Webseite: Webseitensuche	Dokumente	Künstl* - KI - Algorithm* - Learning Analytics	2 (0,1 %)	0 (0,0 %)
HFD	Webseite: (I) Publikationssuche + (II) Themendossier: Künstliche Intelligenz + (III) Themendossier: Learning Analytics	Themendossiers, Blogbeiträge, Webseiteneinträge & Dokumente	Künstl* - KI - Algorithm* - Learning Analytics	30 (2,2 %)	18 (11,2 %)
				1.338 (100,0 %)	161 (100,0 %)

Im ersten Schritt der Dokumentenauswahl wurde die öffentliche Zugänglichkeit als zwingendes Einschlusskriterium geprüft. Insgesamt wurden 40 Dokumente ausgeschlossen, die ausschließlich aus Landtagen stammten und trotz zusätzlicher Recherche nicht abrufbar waren, da es sich etwa um nicht öffentliche Ausschussprotokolle handelte. Daran anschließend wurde die Eignung der Dokumente für die Beantwortung der Forschungsfrage anhand eines Summenindexes dreier dichotomer Kategorien beurteilt (Rel_b – Rel_d).[23] Dafür wurde im zweiten Schritt das Vorkommen relevanter Suchbegriffe erhoben (Rel_b). Da viele Dokumente kein Abstract enthielten, erfolgte diese Prüfung im Volltext. Berücksichtigt wurden die Suchbegriffe der ersten Komponente sowie weitere im Pearl-Growing-Verfahren ergänzte Begriffe.[24] Die starke Variation des Seitenumfangs der Dokumente (Min. = 1 Seite, Max. = 1.299 Seiten) erforderte diese quantitative Vorprüfung. Im dritten Schritt erfolgte eine Prüfung, ob das Dokument Formen politischer Positionierung wie Erklärungen, Argumente, explizite oder implizite Rechtfertigungen oder andere Beiträge zur politischen Debatte über die Rechtfertigung oder Kritik der Einführung von KI im Bildungsbereich berücksichtigt (Rel_c). Schließlich wurde im vierten Schritt bewertet, ob ein KI-Bezug im Bildungskontext vorliegt (Rel_d). Dabei wurde ein weites Begriffsverständnis, angelehnt an Schiff (2022), verwendet, das sowohl konkrete als auch diffuse Thematisierungen von KI und Bildung einschließt: „This includes youth and adult education, training and re-skilling, investing in educational systems, public education and awareness, the need to develop more AI talent (e.g., computer scientists, engineers), and social and ethical issues related to AI and education“ (S. 532).

Auf Basis der drei Kategorien wurde ein Summenscore gebildet und ein Relevanzranking erstellt. Nach Ausschluss von neun Duplikaten erfüllten n = 167 Dokumente alle drei Relevanzkriterien und wurden für die Analyse ausgewählt.[25] Diese wurden anschließend heruntergeladen und archiviert (siehe Anhang B1). Im Verlauf der qualitativen Inhaltsanalyse wurden sechs weitere Dokumente ausgeschlossen, da sie keinen hinreichenden Bildungsbezug aufwiesen (n = 4), lediglich eine Abstimmung darstellten (n = 1) oder sich als Duplikat erwiesen (n = 1). Die finale Stichprobe umfasst n = 161 Dokumente. Davon entfallen 50,3 % auf

[23] Die zu Beginn erhobene Variable zur thematischen Einschlägigkeit des Dokuments (Rel_a) wurde nach dem zweiten Reliabilitätstest einvernehmlich ausgeschlossen, da sie keinen zusätzlichen Erkenntnisgewinn bot. Die thematische Passung ergab sich bereits aus der Auswahl der Institutionen und Suchbegriffe, weshalb in den ersten Codierungsrunden keine Varianz in der Bewertung dieser Variable auftrat.

[24] Die Pearl-Growing-Methode wurde erneut eingesetzt, um zu Beginn der Codierungen keine Dokumente irrtümlicherweise auszuschließen und einen umfassenden Überblick über teils synonym verwendete Begriffe zur Bezeichnung von KI-Anwendungen im Bildungskontext zu erhalten. Das Vorkommen folgender Suchbegriffe wurde codiert: künstlich*, KI, Algo*, learn*, maschin*, neuron*, automat*.

[25] Aus Gründen theoretischer Sättigung wurde nach Abschluss der qualitativen Inhaltsanalyse der n = 167 ausgewählten Dokumente auf eine zusätzliche Codierung der n = 242 Dokumente verzichtet, die in der Auswahl einen Summenscore von zwei Punkten erreichten.

Parlamentsmaterialien des Bundestags und der Landtage, weitere 10,5 % auf Dokumente der politischen Gremien GWK und KMK. Die Beiträge der beratenden Institutionen WR, EFI und TAB bilden zusammen 16,2 % des Gesamtkorpus. Die Stellungnahmen der GEW machen 11,8 % und die des HFD 11,2 % aus. Damit weist der Dokumentenkorpus ein deutliches Übergewicht politischer Positionierungen auf, was in der nachfolgenden Analyse zu berücksichtigen ist. Zugleich sind bildungsspezifische und gewerkschaftliche Perspektiven angesichts der geringen Beteiligung von DL, HLB und HRK unterrepräsentiert. Ein vollständiges Verzeichnis aller Analysedokumente findet sich in Anhang A1.

8.3. Inhaltlich Strukturierende Qualitative Inhaltsanalyse

Auf Basis der ausgewählten Dokumente wird nun angestrebt, die den Rechtfertigungen zugrunde liegenden moralischen Prinzipien durch die Identifikation von Cités in einem explorativen Ansatz sichtbar zu machen. Da Cités dynamische Konstrukte darstellen, die sich weiterentwickeln oder neu formieren können, erfordert ihre Analyse Offenheit. Wie in Abschnitt 5.1 erläutert, besteht das Anliegen der SdK in der „Rekonstruktion der *empirisch* vorhandenen moralischen Ordnungen“ (Diaz-Bone, 2018, S. 145; Herv. i. O.). Ein vollständig standardisiertes Vorgehen erscheint daher wenig geeignet. Stattdessen bietet sich eine qualitative Herangehensweise an, die auch „latente Sinnstrukturen“ (Mayring, 2022, S. 9) erfasst und somit einer potenziellen Weiterentwicklung oder Neuentdeckung der Cités gerecht werden kann.

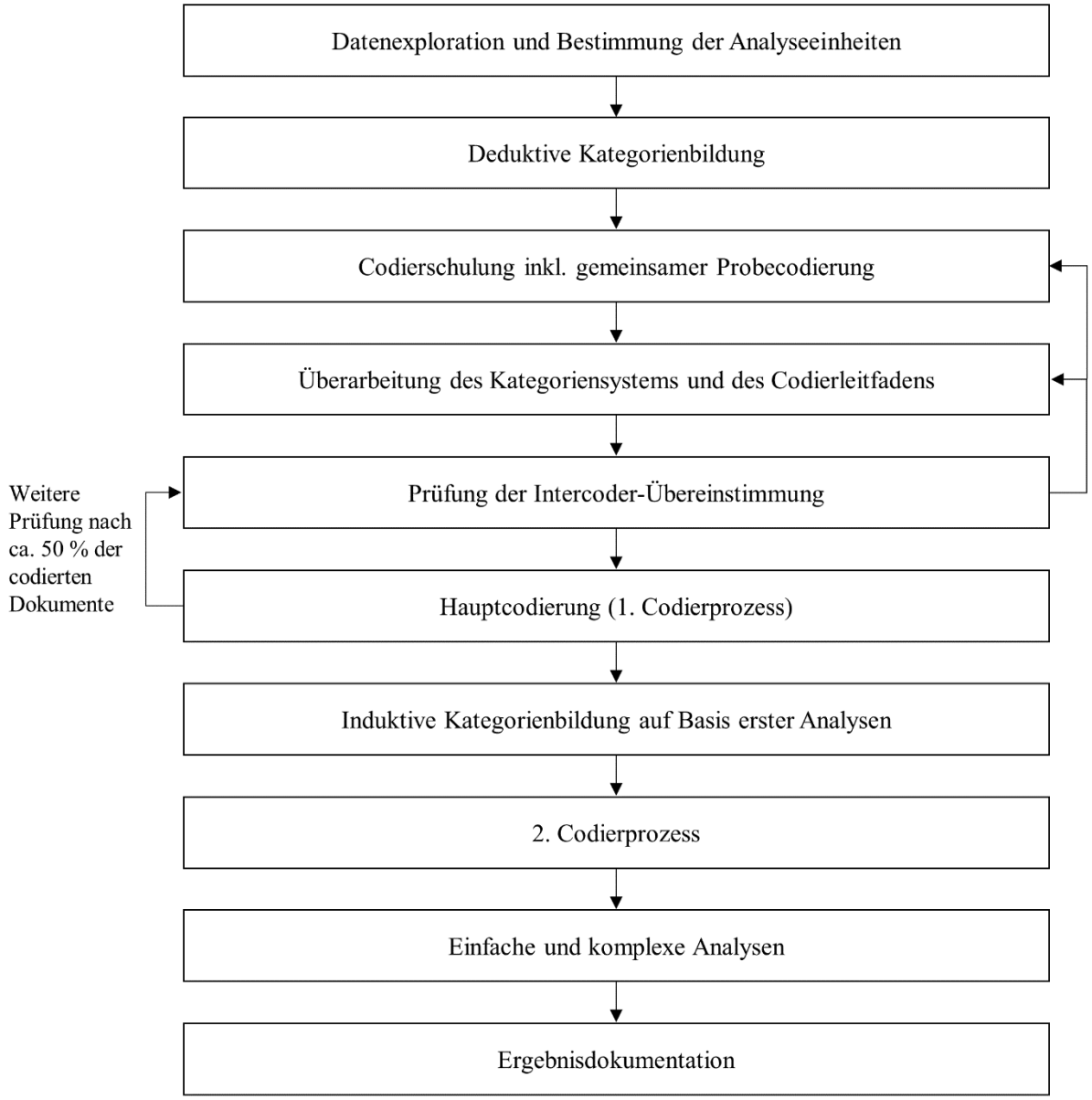
Zur Auswahl eines geeigneten qualitativen Verfahrens wurden verschiedene Ansätze geprüft, darunter Grounded Theory (Glaser & Strauss, 1967/2017; Mey & Mruck, 2011), objektive Hermeneutik (Oevermann et al., 1979), die dokumentarische Methode (Bohnsack et al., 2013a) sowie die qualitative Inhaltsanalyse (Kuckartz & Rädiker, 2022; Mayring, 2022). Dabei zielt die Grounded Theory vor allem auf theoriegenerierende Forschung (Glaser & Strauss, 1967/2017), eignet sich jedoch weniger, wenn wie in diesem Fall ein theoretischer Rahmen bereits vorliegt. Die objektive Hermeneutik fokussiert hingegen stark auf die detaillierte Analyse einzelner Fälle (Wernet, 2006), was in Anbetracht des Ziels einer umfassenden Untersuchung des gesamten Dokumentenkorpus inadäquat erscheint. Besser geeignet erscheinen die dokumentarische Methode und die qualitative Inhaltsanalyse. Erstere rekonstruiert kollektive Orientierungen und handlungsleitende Wissensbestände und setzt auf Fallvergleiche (Bohnsack et al., 2013b). Auch die qualitative Inhaltsanalyse ermöglicht die Untersuchung manifester und latenter Kommunikationsinhalte, geht dabei jedoch stärker regelgeleitet vor. Sie stützt sich auf deduktiv, induktiv oder in der Kombination beider Herangehensweisen entwickelte Kategorien und erlaubt sowohl qualitative als auch quantifizierende Auswertungen (Kuckartz & Rädiker, 2022). Durch ihr systematisches Vorgehen ermöglicht sie zum einen generalisierende Aussagen

über den Untersuchungsgegenstand und zeichnet sich zum anderen durch die Möglichkeit der Bewertung anhand wissenschaftlicher Gütekriterien aus (Mayring, 2022). In Abgrenzung zu evaluativen oder typenbildenden qualitativen Inhaltsanalysen wird für die vorliegende Untersuchung auf den Ansatz einer inhaltlich strukturierenden qualitativen Inhaltsanalyse zurückgegriffen, deren zentrales Ziel in der Identifikation und Systematisierung von Kommunikationsinhalten liegt (Kuckartz & Rädicker, 2022). Dabei wird eine deduktiv-induktive Kategorienbildung gewählt, die es ermöglicht, vorhandenes theoretisches Wissen, etwa über bekannte Cités, in erste Kategorien zu überführen und zugleich offen für neue Erkenntnisse zu bleiben. Quantitativ-statistische Auswertungen ergänzen die Analyse und erlauben eine systematische und transparente Rekonstruktion der Muster in den Rechtfertigungen und der Kritik des KI-Einsatzes im deutschen Bildungssektor. Der zuvor dargelegte theoretische Analyserahmen der techniksoziologischen Deutungsmuster dient der weiteren argumentativen Interpretationsabsicherung der Untersuchung, indem die gewonnenen Erkenntnisse vor diesem Hintergrund reflektiert werden können (Mayring, 2016).

8.3.1. Vorgehen

Sowohl Mayring (2022) als auch Kuckartz und Rädiker (2022) beschreiben die inhaltlich strukturierende qualitative Inhaltsanalyse als mehrschrittigen Prozess. Für die vorliegende Arbeit wurden beide Ansätze zusammengeführt und in eine zehnstufige Auswertungsmethode überführt (siehe Abbildung 3), die nachfolgend dargelegt wird. Der erste Schritt umfasst die Datenexploration und die Festlegung der Analyseeinheiten. Dabei wird zwischen Codier-, Kontext- und Auswertungseinheiten nach Mayring (2022) unterschieden. Die Codiereinheit bildet den kleinsten codierbaren Textbestandteil, hier ein einzelnes Wort (z. B. bei der Codierung von Akteur:innen). Die Kontexteinheit variiert je nach Dokumenttyp, etwa ein Redebeitrag im Plenarprotokoll oder ein ganzer Blogbeitrag. Als Auswertungseinheiten gelten einzelne Sinneinheiten. Um diese zu bestimmen, wurde die initiierende Textarbeit teilautomatisiert durchgeführt. Aus insgesamt 14.927 Dokumentenseiten wurden 15.965 Sätze identifiziert, die mindestens einen der Suchbegriffe der Dokumentenauswahl (Rel_b) enthalten. Diese Sätze dienten als Grundlage für die Festlegung der finalen Sinneinheiten. Dabei wurden im Rahmen der weiteren initiierenden Textarbeit alle automatisch codierten Textstellen durch die drei beteiligten Codierenden auf Relevanz geprüft und bei Bedarf zu vollständigen Sinnabschnitten erweitert (i. d. R. Absatz, nachfolgend als Segment bezeichnet). Als Ergebnis der initiierenden Textarbeit konnten n = 1.067 Segmente identifiziert werden.

Abbildung 3. Vorgehen der inhaltlich strukturierenden qualitativen Inhaltsanalyse



Im zweiten Schritt erfolgte die deduktive Kategorienbildung (Kuckartz & Rädiker, 2022; Mayring, 2022). Auf Grundlage der theoretischen Annahmen wurden Haupt- und erste Nebenkategorien mit Definitionen, Anwendungs- und Abgrenzungsregeln sowie Ankerbeispielen entwickelt und in einem Codierleitfaden dokumentiert (siehe Anhänge A4 und A5). Während das Kategoriensystem im folgenden Abschnitt erläutert wird, erfolgten im dritten Schritt zunächst eine Codierschulung und eine Probecodierung durch die Codierenden. Die Einbindung mehrerer Forschender erhöht die externe Studiengüte der Untersuchung (Denzin, 1970/2017; Lamnek & Krell, 2016), erfordert jedoch besondere Maßnahmen zur Sicherung der Reliabilität. Um eine gemeinsame Anwendung der Kategorien sowie die Regelgeleitetheit der Untersuchung zu gewährleisten (Kuckartz & Rädiker, 2022; Mayring, 2016), wurden mehrere Schulungen durchgeführt, in denen das deduktiv entwickelte Kategoriensystem sowie zentrale Begriffe und die

theoretischen Grundlagen der Cités vorgestellt und in Form des Codierleitfadens bereitgestellt wurden. Im Anschluss codierten in einer ersten Probecodierung alle Codierenden zwei Dokumente unabhängig voneinander, um die Intercoderreliabilität zu berechnen. Die Ergebnisse wurden dem Ansatz des konsensuellen Codierens folgend gemeinsam reflektiert, wobei Nicht-Übereinstimmungen diskutiert und begründet aufgelöst wurden (Guest et al., 2012; Hopf & Schmidt, 1993; Kuckartz & Rädiker, 2022). So wurden auf dieser Grundlage im vierten Schritt der Codierleitfaden und das Kategoriensystem überarbeitet. Im fünften Schritt wurde die Intercoder-Übereinstimmung im Rahmen einer zweiten Probecodierung erneut überprüft. Nach Erreichen einer zufriedenstellenden Übereinstimmung begann im sechsten Schritt die Hauptcodierung. Die Dokumente wurden zufällig den Codierenden zugeteilt und die Segmente entlang der Kategorien codiert. Auch in dieser Phase wurde das konsensuelle Codieren fortgeführt, wodurch laufende Abstimmungen möglich waren und die Präzision der Kategorien weiter erhöht wurde. Entscheidungen, die aus der gemeinsamen Diskussion hervorgingen, wurden zusätzlich im Logbuch dokumentiert. Zur statistischen Sicherung der Reliabilität wurde der Koeffizient Kappa berechnet (Cohen, 1960). Die beiden vor der Hauptcodierung durchgeführten Tests ergaben auf Dokumentebene durchschnittliche Werte von $\kappa = .89$ und $\kappa = .92$. Die insgesamt dritte Prüfung der Intercoder-Übereinstimmung nach circa der Hälfte der codierten Dokumente konnte dieses Ergebnis bestätigen ($\kappa = .93$; siehe Anhang A6), sodass erneut eine sehr gute Intercoderreliabilität gewährleistet wurde (Lombard et al., 2002).

Parallel zur Hauptcodierung wurden darüber hinaus Vorschläge für die induktive Kategorienbildung dokumentiert. Die Bildung datengeleiteter Kategorien im siebten Schritt markiert bereits den Übergang zur Auswertung, da sie eine erste inhaltliche Sichtung und Analyse des Materials erfordert (Kuckartz & Rädiker, 2022). Im Zuge dessen wurden weitere Subkategorien gebildet, definiert und in das Kategoriensystem aufgenommen. Anschließend folgte achtens ein zweiter Codierprozess, um diese Subkategorien präzise zuzuordnen. In den abschließenden Schritten der Analyse erfolgten einfache und komplexe Auswertungen, bei denen Zusammenhänge zwischen Kategorien untersucht wurden. Die Rechtfertigungen wurden dabei sowohl qualitativ als auch quantitativ betrachtet. Die Analyse erfolgte computergestützt mit MAXQDA Analytics Pro 2024 (Version 24.6.0).[26],[27]

[26] Weitere quantitative Auswertungen erfolgten in Excel, da MAXQDA nur Dokumente als Referenzeinheiten erlaubt, während für die Analyse die Segmente als Bezugspunkt dienten. Quantifizierende Aussagen, etwa zu Häufigkeiten und Zusammenhängen zwischen Rechtfertigungen und Kontexten, wurden daher zusätzlich in Excel berechnet (siehe Anhang B3).

[27] Zur sprachlichen Überarbeitung der Arbeit wurden KI-gestützte Werkzeuge eingesetzt. Hierzu zählen ChatGPT (bis Version 5.2) zur stilistischen Überarbeitung einzelner Textpassagen sowie das QuillBot-Word-Add-in (Version 4.37.0) zur Rechtschreib- und Grammatikprüfung. Inhaltliche Verantwortung und wissenschaftliche Bewertung verbleiben vollständig bei der Autorin.

8.3.2. Kategoriensystem

Bei der Erstellung des Kategoriensystems wurden zentrale Qualitätsanforderungen qualitativer Inhaltsanalysen berücksichtigt, insbesondere die Orientierung an den Forschungsfragen, die Präzision der Kategorien sowie die Kohärenz des Gesamtsystems (Kuckartz & Rädiker, 2022). Definitionen, Beispiele und Abgrenzungsregeln im Codierleitfaden unterstützten eine klare Anwendung der Kategorien, während der deduktiv-induktive Ansatz der Kategorienbildung die Vollständigkeit des Systems sicherstellte. Dabei wurden vordefinierte theoriegeleitete Kategorien systematisch verwendet und bei Bedarf durch datengeleitete Kategorien ergänzt. Das vollständige Kategoriensystem inklusive der Codierregeln ist im Anhang A5 dokumentiert. Im Folgenden werden die zentralen Strukturprinzipien und Kategorien erläutert. Zunächst wurden die im Rahmen der Dokumentenauswahl erhobenen formalen Kategorien wie fortlaufende Nummer, Institution, Dokumententitel, Dokumententyp und Erscheinungsdatum als Dokumentenvariablen in MAXQDA importiert. Das inhaltliche Kategoriensystem umfasst drei Ordnungskategorien, die erstens der Strukturierung dienen (Block A), zweitens die Kategorien des KI-Kontextes zusammenfassen (Block B) und drittens Kategorien der Rechtfertigung umfassen (Block C). Zur Strukturierung der Analyse wurden die automatisiert identifizierten Sätze, die mindestens einen Suchbegriff enthalten (A.1), sowie die daraus abgeleiteten Segmente (A.2) in zwei Faktenkategorien dargestellt. Die Kategorie zitierfähiger Aussagen (A.3) ermöglicht es, besonders prägnante Passagen für die Ergebnisdarstellung zu markieren. Unklarheiten (A.4) sowie neue Ideen (A.5) wurden ebenfalls codiert, um das konsensuelle Vorgehen bei Unsicherheiten zu unterstützen und Vorschläge für induktiv zu ergänzende Kategorien systematisch zu erfassen.

Der Kontext (Block B) erfasst den Anwendungsbereich des KI-Bezugs (B.1) sowie den Bildungsbereich (B.2). Im Anwendungsbereich wird zwischen konkreten (B.1.1) und diffusen KI-Bezügen (B.1.2) unterschieden. Konkrete KI-Anwendungen (B.1.1.1 – B.1.1.6) wurden in Anlehnung an die durch Chiu et al. (2023) vorgenommene Systematisierung von KI-Anwendungen im Bildungsbereich den Bereichen Lernen, Lehren, Bewertung, Forschung, Administration oder Sonstiges zugeordnet. Ergänzend dazu wurde die Unterscheidung von Schiff (2022) zwischen dem Einsatz von KI *in* der Bildung und der Bildung *für* den KI-Einsatz berücksichtigt. Vor diesem Hintergrund wurden die Kategorien Wissenserwerb über KI, Investitionen in KI sowie Forschung über KI und sonstige KI-Kontexte (B.1.3 – B.1.6) aufgenommen, um sowohl bildungspolitische Strategien zur Vorbereitung einer KI-fähigen Gesellschaft als auch den Einsatz von KI-Systemen im Bildungsbereich systematisch abzubilden. Die Verortung im Bildungssystem erfolgte darüber hinaus entlang der Strukturbeschreibung des deutschen

Bildungssystems der KMK (2021). Codiert wurden der Elementarbereich, der frühkindliche Bildungseinrichtungen umfasst, der Primar- und Sekundarbereich mit allen Schulformen von der Grundschule bis zur Sekundarstufe II, der Tertiärbereich, der die Hochschulbildung und -forschung einschließt, sowie der Quartärbereich (B.2.1 – B.2.4), der Weiterbildung und Erwachsenenbildung umfasst. Eine Ausdifferenzierung zwischen Primar- und Sekundarbereich wurde im Zuge der Analyse zugunsten einer gemeinsamen Kategorie aufgegeben, da die politische Debatte beide Bereiche meist unter dem Sammelbegriff „Schule" zusammenfasst. Zudem wurde im Zuge der induktiven Kategorienbildung die Kategorie des unspezifizierten Bildungsbereichs (B.2.5) ergänzt, um Fälle abzudecken, in denen keine eindeutige Zuordnung möglich war.

Block C umfasst mit den Rechtfertigungen den Kern der Analyse. Erfasst werden die Cité (C.1), die Form der Prüfung (C.2), die Richtung der Rechtfertigung (C.3) sowie rechtfertigende Akteur:innen (C.4). Die Subkategorien der Cités entsprechen den neun in Abschnitt 5.1 dargestellten Cités (C.1.1 – C.1.9): die Cité der Inspiration, des Hauses, der Meinung, des Staatsbürgers, der Industrie, des Marktes, des Projekts, der Ökologie sowie der Solution. Ergänzt wurden Kategorien für andere Formen der Rechtfertigung (C.1.10), um im Sinne der Offenheit die Möglichkeit zu wahren, neu herausgebildete Rechtfertigungsmuster identifizieren zu können, sowie mangelnde Rechtfertigung (C.1.11), da im Zuge der Analyse vermehrt ungerechtfertigte Äußerungen identifiziert wurden. Außerdem konnten bereits einige Konfliktlinien zwischen verschiedenen Cités aufgedeckt werden. Um daher die Tiefe der bildungspolitischen Auseinandersetzung über die Einführung von KI-Systemen nicht nur qualitativ, sondern auch quantitativ beschreiben zu können, wurden die analytischen Kategorien der Wahrheits-, der Realitäts- und der Existenzprüfung sowie des Kompromisses (C.2.1 – C.2.4) in das Kategoriensystem aufgenommen und im zweiten Codierprozess erhoben. Die Richtung der Rechtfertigung wurde über evaluative Kategorien abgebildet (C.3.1 – C.3.5), die unterscheiden, ob eine Aussage für, abwägend positiv, neutral, abwägend negativ oder gegen den KI-Einsatz beziehungsweise eine Beschäftigung mit KI (in Form von Wissenserwerb, Investitionen oder Forschung) ausfällt. Schließlich unterscheiden die Subkategorien der rechtfertigenden Akteur:innen (C.4) politische (C.4.1), wissenschaftliche (C.4.2), gewerkschaftliche (C.4.3), verbandspolitische[28] (C.4.4) und wirtschaftliche Akteur:innen (C.4.5). Politische Akteur:innen wurden, sofern erkennbar, nach der Parteizugehörigkeit aller im Untersuchungszeitraum vertretenen Bundestagsfraktionen

[28] Äußerungen (von Vertreter:innen) des HFDs wurden als „Verband" codiert im Sinne eines Zusammenschlusses gemeinsamer Interessen und Ziele (Schubert & Klein, 2021).

(C.4.1.1.1 – C.4.1.1.7) sowie nach Regierungen[29] (C.4.1.2) oder parlamentarischen Gremien (C.4.1.3), etwa der Enquete-Kommission, differenziert. Wissenschaftliche Akteur:innen umfassen wissenschaftliche Beratungsgremien wie TAB, WR und EFI (C.4.2.1 – C.4.2.3) sowie einzelne Wissenschaftler:innen verschiedener Disziplinen (Sonstige: C.4.2.4). Diese Kategorien dienen als Grundlage der nachfolgenden Ergebnisdarstellung.

9. Ergebnisse

Die nachfolgende Analyse basiert auf der qualitativen und quantitativen Auswertung der insgesamt n = 1.067 Segmente, die jeweils als eine Sinneinheit in n = 161 Dokumenten definiert sind. Quantifizierend werden die Häufigkeitsverteilungen analyserelevanter Kategorien auf zwei Ebenen betrachtet: zum einen relational zu den codierten Segmenten, zum anderen relational zur Gesamtzahl der Cité-Bezüge. Darüber hinaus werden qualitative Einblicke in die inhaltlichen Schwerpunkte sowie die moralischen Grundsätze und Konfliktlinien geboten. Die in diesem Zusammenhang vermerkten Verweise auf Dokumente und Zitate sind dabei exemplarisch und nicht erschöpfend zu verstehen.

Für die Strukturierung der Ergebnisdarstellung bieten sich grundsätzlich eine thematische oder eine analytische Gliederung an. Erstere bündelt zentrale Schwerpunkte der Debatte, würde jedoch die Rekonstruktion der Cités über die Darstellung hinweg zerstreuen und zugleich den Bildungssektor als exemplarischen Anwendungsfall übermäßig in den Vordergrund rücken. Eine analytische Gliederung ermöglicht hingegen eine systematische Rekonstruktion der Rechtfertigung und Kritik des KI-Einsatzes entlang der Forschungsfragen und macht damit verbundene intra- und interzitäre Spannungen sichtbar. Sie führt zwar dazu, dass einzelne Themen an verschiedenen Stellen erneut aufgegriffen werden, jedoch jeweils aus der Perspektive unterschiedlicher Cités. Dennoch wird nachfolgend die analytische Struktur gewählt, da sie eine präzisere Herausarbeitung konventionslogischer Muster erlaubt und die Übertragbarkeit der Befunde über den untersuchten Fall hinaus verdeutlicht.

Im ersten Schritt wird der bildungspolitische Kontext der KI-Debatte umrissen. Dabei wird aufgezeigt, welche bildungspolitischen Akteur:innen worüber sprechen, wenn von „KI“ im Bildungskontext die Rede ist. Im zweiten Schritt folgt ein quantifizierender Überblick über die

[29] Erfasst wurden Äußerungen im Namen von Bundes- oder Landesregierungen, sofern sie ohne parteispezifischen Bezug erfolgten. Es wurde keine explizite Unterscheidung zwischen Regierung und Opposition vorgenommen, stattdessen wurden oppositionelle Positionen über die Parteizugehörigkeit abgebildet. Eine weitergehende Differenzierung erschien forschungsökonomisch nicht sinnvoll, da akteursspezifische Rechtfertigungsmuster nicht im Zentrum der Untersuchung stehen und eine dezidierte Trennung angesichts wechselnder Regierungs- und Oppositionskonstellationen über acht Jahre auf Bundes- und Landesebene einen unverhältnismäßigen Mehraufwand bedeutet.

Verteilung der Cités, die Valenz der Debatte sowie das Vorkommen der Prüfungsformen der SdK. Diese Ergebnisse werden im dritten Schritt qualitativ vertieft. Zunächst werden die Rechtfertigungen des KI-Einsatzes analysiert, um zu klären, welche moralischen Grundlagen und Motive den Forderungen nach dem Einsatz von KI im Bildungswesen zugrunde liegen und wie Legitimation erzeugt werden soll (FF1). Dabei werden bestehende Kompromisse sichtbar und die besondere Rolle der Cité der Solution für die KI-Rechtfertigung herausgearbeitet (FF2). Anschließend wird die Kritik des KI-Einsatzes dargelegt, einschließlich der damit verbundenen Konfliktlinien und Prüfungsformen (FF3). Die Ergebnisse legen offen, welche Spannungen die Debatte prägen, wie Akteur:innen darauf reagieren und wie ausgeprägt die Politisierung der Debatte ist. Diese schrittweise Vorgehensweise ermöglicht ein differenziertes Bild des bildungspolitischen KI-Diskurses, das die vielfältigen Kontexte, die moralischen Rechtfertigungen und die zugrunde liegenden Konflikte sichtbar macht.

9.1. Kontexte der bildungspolitischen Debatte über Künstliche Intelligenz

KI erscheint im bildungspolitischen Diskurs häufig als vage definierter, vieldeutiger Sammelbegriff, der ein breites Spektrum an Funktionen, Methoden und Anwendungsfeldern umfasst. Um zu klären, was konkret mit „KI“ in bildungspolitischen Zusammenhängen assoziiert wird, wurde in einem ersten Analyseschritt eine Kontextualisierung vorgenommen. In den Blick genommen wurden der Anwendungsbereich des KI-Bezugs sowie die adressierten Bildungsbereiche. Erstere umfassen drei Arten von KI-Bezügen: (1) Bezüge auf konkrete KI-Anwendungen, die definierbaren Anwendungsbereichen in den Feldern Lernen, Lehre, Bewertung von Lernenden, Administration oder Forschung zugeordnet werden können und im folgenden Abschnitt differenzierter betrachtet werden. (2) Diffuse Verweise auf KI, bei denen kein spezifischer Anwendungskontext erkennbar ist. (3) Bezüge, die sich nicht auf Anwendungen selbst, sondern auf den Umgang mit KI richten, in Form des Wissenserwerbs oder der Forschung über KI sowie Investitionen in KI. Die Analyse zeigt, dass verschiedene KI-Bezüge häufig verbunden werden, sodass Mehrfachnennungen innerhalb eines Segments codiert wurden.

Konkrete KI-Anwendungen werden mit einem Anteil von 58,9 % aller Segmente am häufigsten thematisiert. Jeweils rund ein Fünftel der Segmente behandelt diffuse KI-Bezüge ohne klare Verortung (20,0 %), den Wissenserwerb über KI (19,5 %) oder Investitionen in KI (20,8 %). Etwas seltener wird Forschung über KI (16,6 %) adressiert. Mit Blick auf die adressierten Bildungsstufen fokussiert sich die KI-Debatte deutlich auf die hochschulische Bildung und Forschung, das heißt auf den tertiären Bereich (48,3 %), der in nahezu der Hälfte aller Segmente als Referenzpunkt dient. Zugleich bleibt der Bildungsbereich in vielen Segmenten unspezifiziert (36,3 %). Der Elementarbereich wird in 4,5 %, der Primar- und Sekundarbereich in 16,8 %

und der Quartärbereich in 8,5 % der Segmente zum Gegenstand der KI-Debatte. Im Folgenden werden die inhaltlichen Bezüge zu KI anhand von Beispielen schematisch entlang der identifizierten Anwendungs- und Diskursbereiche entfaltet. Ziel ist es, die konkrete Bedeutungszuschreibung von „KI“ im Bildungskontext sichtbar zu machen und damit die an späterer Stelle dargelegten Rechtfertigungsmuster der bildungspolitischen Debatte zu konturieren.

9.1.1. Differenzierung der thematisierten Anwendungen Künstlicher Intelligenz

Mit 58,9 % der Segmente prägt der Bezug auf tatsächliche KI-Anwendungen maßgeblich die bildungspolitische Debatte. Am häufigsten werden Anwendungen im Bereich des Lernens thematisiert, die in 35,2 % der Segmente auftreten. Seltener finden sich Verweise auf KI-Anwendungen in der Lehre (11,7 %), zur Bewertung von Lernenden (6,3 %), in der Verwaltung (11,9 %) sowie auf die forschungsbezogene Nutzung von KI-Systemen (7,6 %). Während die exemplarischen Anwendungen der genannten Bereiche an dieser Stelle nicht ausführlich besprochen werden können (für eine Übersicht siehe Tabelle 6), sollen zwei zentrale Befunde hervorgehoben werden, die sich bereichsübergreifend zeigen.

Erstens zeigt sich, dass die Grenze zwischen Digitalisierung und KI in vielen Fällen unscharf bleibt. Dies betrifft insbesondere den Bereich des Lernens, in dem sowohl Anwendungen, die im Sinne der in Kapitel 2 dargelegten Definition als KI zu verstehen sind (21,8 %), als auch einfache digitale Lösungen (13,4 %) unter dem KI-Begriff subsumiert werden. Dazu zählen vor allem Blended- und E-Learning-Formate, die in unmittelbare Nähe zum KI-Begriff gerückt werden (Brandenburg/FDP, BT_199c, 2019, S. 13133; BT_406b, 2021, S. 5; LT-TH_733, 2020, S. 6; Weltzien/Die Linke, LT-TH_736, 2020, S. 1520; Günther/CDU, LT-SH_830, 2021, S. 8742; LT-SH_934, 2021, S. 23, 33), wodurch der Eindruck entsteht, Digitalisierung und KI seien weitgehend synonym zu verstehen und zu behandeln. Mitunter werden Blended-Learning-Ansätze auch als Voraussetzung des Erlernens des Umgangs mit KI-Anwendungen verstanden (BT_113b, 2020, S. 340). Hinzu kommt, dass teilweise die Grenzen zwischen einzelnen Anwendungsbereichen verschwimmen, wenn betont wird, dass die jeweiligen Systeme Lernenden und Lehrenden gleichermaßen zugutekommen sollen (Völlers/SPD, BT_244c, 2019, S. 14626; HFD_1190, 2016).

Zweitens zeigt die Analyse, dass konkrete Anwendungsbeispiele nur selten benannt werden. Zu den wenigen Ausnahmen gehören Telepräsenz-Avatare, die erkrankten Schüler:innen die virtuelle Teilnahme am Unterricht ermöglichen sollen (LT-HE_979, 2022, S. 1), sowie das adaptive Lernsystem „area9 Rhapsode“, das Lernparameter (z. B. Wissenslücken, Selbstvertrauen, Bearbeitungszeiten) analysiert und individuelle Lernpfade erstellt (LT-SN_894, 2021, S. 2 f.).

Tabelle 6. Übersicht thematisierter KI-Systeme nach Anwendungsbereichen

Bereich	Diskutierte KI-Anwendungen	Exemplarische Referenzen
Lernen (35,2 %)	(i) *Analyse von Lernverhalten und Leistung:* (a) Empfehlungs- und Tutorensysteme, z. B. „area9 Rhapsode“, (b) Learning Analytics	BT_244a, 2019; BT_288a, 2020, S. 2; LT-SN_894, 2021
	(ii) *Anpassung und Gestaltung der Lernumgebung:* (a) Blended Learning/E-Learning, (b) Personalisierte Lernnetzwerke, (c) Virtual Reality sowie vereinzelt (d) Telepräsenz-Avatare und (e) Übersetzung und Transkription von Lehrmaterialien	BT_199c, 2019, S. 13133; BT_113b, 2020, S. 298 f., 340; GWK_1052, 2020; HFD_1202, 2021, S. 49 ff.; HFD_1196, 2018, S. 2; LT-HE_979, 2022; LT-NI_660, 2019, S. 3; TAB_1335, 2020, S. 14;
Lehre (11,7 %)	(i) *Entlastung von Routineaufgaben:* z. B. Vokabelabfrage, Hausaufgaben- oder Anwesenheitskontrolle	BT_244b, 2020, S. 5 f.
	(ii) *Beobachtung & Evaluation der Lehre:* (a) virtuelle Schulbücher, (b) Eye-Tracking-Anwendungen, (c) Lernzielauswertung auf Modul- oder Klassenebene, (d) Qualitätsentwicklung durch „KI-Teaching-Assistant“	BT_113b, 2020, S. 333 ff.; BT_288a, 2020, S. 2 f.; HFD_1201, 2021, S. 2; LT-SN_894, 2021, S. 2
	(iii) *Lehr-Gestaltung und -Weiterentwicklung:* Lehrangebote, Kurse, Module, Curricula	GWK_1052, 2020, S. 2; HFD_1202, 2021, S. 33 ff.; HFD_1206, 2021, S. 2
	(iv) *Immersive Lehrformate:* Virtual Reality/virtuelle Labore z. B. für naturwissenschaftliche Experimente oder historische Ereignisse	BT_113b, 2020, S. 298 f.; S. 338 f.; HFD_1202, 2021, S. 31; LT-NW_895, 2021, S. 89; WR_1177, 2022, S. 94
Bewertung (6,3 %)	(i) *Bewertung & Benotung von Prüfungsleistungen:* Multiple-Choice- und Freitextklausuren	HFD_1187, 2015, S. 18, 51; HFD_1210, 2021, S. 1 ff.; LT-HE_765, 2020, S. 6; TAB_1335, 2020, S. 66 f.
	(ii) Prognose von Studienabbrüchen	BT_113b, 2020, S. 338 f.; HFD_1202, 2021, S. 27 ff.; TAB_1333, 2021, S. 4; WR_1177, 2022, S. 38
	(iii) Verhaltensbewertungssysteme (vereinzelte Nennungen), z. B. ClassDojo aus den USA	GEW_1244, 2019, S. 11
Administration (11,9 %)	*(i) Studienberatung:* insbesondere durch Chatbots	HFD_1201, 2021, S. 1; GWK_1052, 2020, S. 2; LT-SH_970, 2022, S. 10724; WR_1177, 2022, S. 39
	(ii) Personalprozesse: (a) Lehrkräftepersonalverwaltung, (b) Bewerber:innenauswahl, (c) Fort- & Weiterbildungsempfehlungen	BT_113b, 2020, S. 197 f.; LT-SH_670, 2019, S. 35
	(iii) Antragsprozesse: z. B. Bearbeitung von BAföG-Anträgen	BT_455c, 2022, S. 29
Forschung (7,6 %)	(i) *Forschungsinfrastruktur:* Bedarf an (a) Hochleistungsrechnern, (b) Wissens- & (c) Publikationsplattformen	EFI_1082, 2019, S. 96; GWK_1060, 2022, S. 131 f.
	(ii) Analyse & Simulation in Forschungsprozessen: (a) Datenanalyse, (b) Analyse wissenschaftlicher Publikationen	BT_113b, 2020, S. 299 f.; GWK_1044, 2018, S. 395

Anmerkung: Die Prozentangaben der einzelnen Bereiche geben den Anteil des Vorkommens eines jeden Bereichs in den codierten (n = 1.067) Segmenten an. Mehrfachnennungen möglich.

Häufig bleibt die Darstellung hingegen vage. So wird der Einsatz von „KI“ oder „Learning Analytics“ zur Individualisierung von Lernprozessen gefordert, ohne dass ersichtlich wird, welche Verfahren oder Anwendungen konkret eingesetzt werden sollen oder wie diese zur erwarteten Verbesserung beitragen könnten (Die Linke, BT_195b, 2019, S. 10; BT_406b, 2021, S. 5; KMK_1076, 2021, S. 14; Wulf/CDU, LT-NI_664, 2019, S. 5799; Voigt/CDU, LT-TH_850, 2021, S. 3551; LT-HB_900, 2021, S. 2; Promny/FDP, LT-HE_1011, 2022, S. 8584). Ähnliche Muster zeigen sich in kritischen Beiträgen, die KI problematisieren, jedoch ohne Bezug zu konkreten Anwendungsszenarien (Espendiller/AfD, BT_244c, 2019, S. 14548; Jankowski/AfD, LT-TH_850, 2021, S. 3550).

Diese Befunde verweisen darauf, dass der Bezug auf KI-Anwendungen in verschiedenen Einsatzbereichen zwar den Ausgangspunkt der Debatte bildet, diese aber häufig nur oberflächlich oder unscharf beschrieben werden. Dies wird zusätzlich durch diffuse KI-Bezüge verstärkt, die keinen klaren Anwendungsbereich adressieren und im Folgenden dargestellt werden.

9.1.2. Diffuse Vorstellungen des Einsatzes Künstlicher Intelligenz im Bildungswesen

Die genannten Befunde setzen sich in den Ergebnissen zu diffusen KI-Anwendungsbezügen fort, die in 20,0 % der analysierten Segmente nachzuweisen sind. Die Analyse bestätigt eine breite, häufig unspezifizierte Verwendung des KI-Begriffs, die die von Köstler und Ossewaarde (2022) beschriebene „Verwässerung“ des Begriffs auch im bildungspolitischen Diskurs sichtbar macht. Häufig wird betont, dass die Chancen des KI-Einsatzes genutzt werden müssten (BT_162a, 2018, S. 1; BT_162b, 2019, S. 2; Brandenburg/FDP, BT_199c, 2019, S. 57, S. 13133; Bundesregierung, BT_244b, 2020, S. 9; Suding/FDP, BT_244c, 2019, S. 14547; Karliczek/Union, BT_315c, 2020, S. 21845; BT_406a, 2021, S. 2; GWK_1058, 2021, S. 1; HFD_1188, 2016, S. 1; LT-SH_787, 2020, S. 4; Promny/FDP, LT-HE_1011, 2022, S. 8583; Lorz/CDU, LT-HE_1011, 2022, S. 8594; WR_1177, 2022, S. 49). Dabei bleiben jedoch technische Funktionsweisen oder konkrete Nutzungsszenarien meist offen. KI erscheint häufig als abstrakte Ressource, die Lernprozesse oder Bildungsqualität verbessern soll, ohne dass diese Wirkungen durch konkrete Anwendungsbeispiele begründet werden. So formuliert die Bund-Länder-Vereinbarung „Künstliche Intelligenz in der Hochschulbildung“[30] das Ziel einer Verbesserung der Hochschulbildung durch KI, wobei der KI-Einsatz weitgehend unspezifisch bleibt (GWK_1052, 2020, S. 1). Ebenso bleibt die im niedersächsischen Landtag geführte

[30] Die Bund-Länder-Vereinbarung „Künstliche Intelligenz in der Hochschulbildung“ ist im Wortlaut neben der Originärpublikation der GWK (GWK_1052, 2020) Gegenstand weiterer Dokumente (GWK_1061, 2022; LT-BE_774, 2020; LT-SH_787, 2020). Um Mehrfachcodierungen und eine Übergewichtung zu vermeiden, wurde ausschließlich die Originärpublikation codiert. Da es sich jedoch nicht um vollständige Duplikate handelt, wurden in den übrigen Dokumenten weitere Segmente identifiziert und codiert.

Debatte über den Einfluss von KI auf den Bildungsbereich ohne Bezug auf konkrete Anwendungsfelder (LT-NI_547, 2019, S. 3913). Des Weiteren zeigt sich die Diffusität auch in pauschalen Verweisen auf ein breites Spektrum möglicher KI-Einsatzszenarien (GWK_1045, 2019, S. 111; HFD_1194, 2017, S. 2; HFD_1206, 2021, S. 1; Alt/SPD, LT-RP_790, 2020, S. 30; Plambeck/CDU, LT-SH_830, 2021, S. 8743) sowie im Kontext von Forderungen nach Regulierung oder gemeinwohlorientierter KI-Gestaltung (BT_162b, 2019, S. 2; Sitte/Die Linke, BT_113d, 2020, S. 23760; LT-RP_522, 2019, S. 1; LT-SH_934, 2021, S. 24).

Insgesamt verdeutlichen diese Befunde, dass ein erheblicher Teil der Debatte auf generalisierten oder unscharfen Vorstellungen von KI basiert. Ergänzend wird im Folgenden dargestellt, welche Rolle die Thematisierung des Wissenserwerbs über KI in der politischen Debatte spielt.

9.1.3. Wissenserwerb über Künstliche Intelligenz

Neben konkreten oder diffusen Bezügen auf KI-Anwendungen bildet der Wissenserwerb über KI einen weiteren zentralen Bestandteil der bildungspolitischen Debatte. In den analysierten Segmenten werden Kompetenzen von der frühkindlichen Bildung bis zur Erwachsenenbildung diskutiert. Die Thematisierung des Wissenserwerbs gliedert sich entlang vier Bezugsgruppen:

(1) *KI-Kompetenzen von Lernenden*: Wiederkehrend wird die Notwendigkeit betont, Medien-, Digital- und KI-Kompetenzen als festen Bestandteil der allgemeinen Bildung zu verankern. Lernende aller Bildungsstufen und Fachrichtungen sollen grundlegende Kenntnisse darüber erwerben, wie digitale Medien und Algorithmen funktionieren (Meyer/CDU, LT-SN_904, 2021, S. 2786; GEW_1218, 2017, S. 2; LT-SH_934, 2021, S. 13). Neben technischem Wissen werden ebenfalls ethische und gesellschaftliche Fragestellungen hervorgehoben, etwa der verantwortungsvolle Umgang mit digitalen Technologien (LT-NI_666, 2019, S. 2 ff.; WR_1175, 2022, S. 54 f.).

(2) *Aus- und Weiterbildung von Lehrkräften:* Lehrkräfte werden als zentrale Multiplikator:innen für den Erwerb von KI-Kompetenzen betrachtet. Gefordert werden sowohl technisches als auch medienpädagogisches und didaktisches Wissen, um KI-Inhalte fundiert und praxisnah vermitteln zu können (BT_113b, 2020, S. 340 f.; BT_288a, 2020, S. 3; HFD_1202, 2021, S. 52 ff.). Neben der Integration von KI-Inhalten in die Lehramtsausbildung wird ein kontinuierlicher Qualifizierungsbedarf hervorgehoben, um auf die Dynamik technologischer Entwicklungen vorbereitet zu sein (TAB_1334, 2021, S. 15). Suding (FDP) fordert etwa, Lehrkräfte müssten frühzeitig lernen, „wie sie ihren Unterricht mithilfe von Daten verbessern können“ (BT_244c, 2019, S. 14547).

(3) *Qualifikation von Fachkräften:* Die Qualifikation und Weiterbildung von Erwerbstätigen und Fachkräften wird als zentrale Voraussetzung beruflicher Teilhabe und

Innovationsfähigkeit betrachtet (FDP, BT_288a, 2020, S. 4). Dabei sollen Weiterbildungsplattformen den Zugang zu KI-Kursen erleichtern und digitale Kompetenzen fördern (BT_113b, 2020, S. 341 ff.; GWK_1060, 2022, S. 105; TAB_1335, 2020, S. 251). Zudem werden die Integration von KI-Inhalten in akademische Curricula sowie die Einrichtung neuer Studiengänge thematisiert, um einem Fachkräftemangel zu begegnen, ergänzt durch KI-bezogene Berufungen an Hochschulen (BT_157a, 2018, S. 4; GWK_1052, 2020, S. 1 f.; HFD_1202, 2021, S. 33 ff.).

(4) *Gesellschaftliche Bildung:* Schließlich wird ein niedrigschwelliges Bildungsangebot für alle Bürger:innen gefordert, etwa über Volkshochschulen. Es umfasst ebenfalls grundlegende digitale Kompetenzen als Voraussetzung für eine selbstbestimmte Nutzung von KI und gesellschaftliche Teilhabe (BT_288b, 2021, S. 12; GEW_1275, 2021, S. 2).

Diese Ergebnisse zeigen, dass der Wissenserwerb über KI im Diskurs nicht nur als notwendige Voraussetzung für den kompetenten Umgang mit technologischen Entwicklungen verstanden wird, sondern auch als zentrale Bedingung für eine verantwortungsvolle Gestaltung von Bildungs- und Arbeitsprozessen.

9.1.4. Forschung über Künstliche Intelligenz

Die Diskussion über Forschung im Bereich der KI bildet einen weiteren Bestandteil der bildungspolitischen KI-Debatte. Dabei lassen sich entlang der codierten Segmente drei Schwerpunkte differenzieren:

(1) *Interdisziplinäre und internationale Forschungskooperationen:* Insbesondere der Bedarf an Interdisziplinarität und internationaler Vernetzung ist ein breit diskutiertes Thema im Rahmen des Forschungsschwerpunkts (BT_113b, 2020, S. 350; GWK_1047, 2019, S. 240; LT-NW_753, 2020, S. 34 f.; LT-NW_895, 2021, S. 165 f.). Um den wissenschaftlichen Austausch zu fördern, werden die Einrichtung neuer Professuren und interdisziplinärer Graduiertenkollegs gefordert (GWK_1051, 2020, S. 396; WR_1181, 2022, S. 43) sowie der Ausbau internationaler und regionaler Kompetenzzentren, etwa im Rahmen des Cyber Valley und europäischer Innovationscluster (BT_157b, 2019, S. 10).

(2) *Anwendungsorientierte Forschung und Wissenstransfer:* Die Enquete-Kommission hebt hervor, dass erfolgreiche KI-Forschung im Bereich der Bildung eine enge Verzahnung von (problemorientierter) Grundlagenforschung und anwendungsorientierter Forschung erfordere (BT_113b, 2020, S. 288). Im Fokus steht somit ebenso der Transfer zwischen Wissenschaft und Praxis (BT_113a, 2018, S. 3 f.; BT_157b, 2019, S. 10; Karliczek/CDU, BT_157d, 2019, S. 9472; GWK_1047, 2019, S. 396).

(3) *Ethische und gesellschaftliche Forschung:* Neben technischen und anwendungsbezogenen Fragen wird die Erforschung ethischer Dimensionen von KI betont. Diskutiert werden sowohl ethische Anforderungen an KI-Forschungsprojekte selbst (GWK_1060, 2022, S. 247) als auch die Notwendigkeit einer ethisch verantwortlichen Entwicklung von KI-Anwendungen (HFD_1211, 2021, S. 8 f.; LT-NW_895, 2021, S. 174).

Die Ergebnisse der Thematisierung von Forschung über KI zeigen damit, dass diese im Bildungsbereich sowohl technisch-wissenschaftliche als auch ethisch-gesellschaftliche Dimensionen umfasst und stark auf Interdisziplinarität, Anwendungsorientierung und Vernetzung ausgerichtet ist.

9.1.5. Investitionen in Künstliche Intelligenz

Die Thematisierung von Investitionen in KI, die in 20,8 % der analysierten Segmente nachzuweisen ist, weist Schnittmengen zu den Debatten über Wissensweb und Forschung über KI sowie zu den Diskussionen um konkrete und diffuse KI-Anwendungen auf. Im Zentrum stehen staatliche Maßnahmen, die den Ausbau von Infrastruktur, Forschungskapazitäten und Qualifizierungsangeboten unterstützen sollen. Drei Bereiche lassen sich unterscheiden:

(1) *Investitionen in Forschung:* Ein wesentlicher Aspekt der Investitionspolitik befasst sich mit der Stärkung der KI-Forschung, darunter die Einrichtung von 100 KI-Professuren (BT_157b, 2019, S. 12; BT_162a, 2018, S. 4) sowie der Auf- und Ausbau nationaler und regionaler KI-Kompetenzzentren (BT_162b, 2019, S. 3 ff.; LT-HE_765, 2020, S. 46). Diese Investitionen werden zugleich als Beitrag zur Fachkräfteentwicklung verstanden (BT_162b, 2019, S. 20; GWK_1052, 2020, S. 1).

(2) *Investitionen in Kompetenzentwicklung:* Gefordert werden ebenso Investitionen in Weiterbildungsangebote, um einen kontinuierlichen Kompetenzaufbau in schulischen und beruflichen Kontexten zu ermöglichen (BT_244d, 2021, S. 59; LT-NW_895, 2021, S. 74; LT-TH_836, 2021, S. 8 f.). Dazu zählen auch Mittel für die (Weiter-)Entwicklung von Studiengängen (LT-NW_895, 2021, S. 165) sowie für die Qualifizierung von Lehrkräften und Studierenden (BT_195b, 2019, S. 2).

(3) *Investitionen in Infrastruktur und Anwendungen:* Ein weiterer Schwerpunkt liegt auf der Modernisierung digitaler Infrastruktur (LT-SH_970, 2022, S. 55 ff.). Ein Beispiel bildet der „DigitalPakt Schule“ des Bundes mit einem Volumen von 5 Milliarden Euro, um Glasfaseranschlüsse und IT-Grundstrukturen zu schaffen, die einen künftigen KI-Einsatz ermöglichen sollen (BT_113b, 2020, S. 337). Auch der Ausbau leistungsfähiger Rechenzentren wird als Voraussetzung für den Einsatz von KI hervorgehoben (BT_157a, 2018, S. 4; Meyer/CDU, LT-SN_904, 2021, S. 2792 f.). Ergänzend werden

> Investitionen in konkrete KI-Anwendungen im Verwaltungs- sowie im Lehr-Lern-Kontext angesprochen (LT-SH_592, 2019, S. 5 f.; LT-HE_765, 2020, S. 43; LT-NW_895, 2021, S. 89), etwa durch landesspezifische Fördermittel wie das Sondervermögen in Schleswig-Holstein zur Finanzierung von KI-Maßnahmen (LT-SH_670, 2019, S. 55).

Die Analyse der Kontextualisierung von KI zeigt damit ein breites Spektrum an inhaltlichen Bezugspunkten, das von Vorstellungen über Einsatzmöglichkeiten bis hin zu Fragen des Wissensaufbaus, der Forschung und der Investitionspolitik reicht, dabei jedoch insbesondere in der Diskussion der Anwendungen häufig diffus bleibt. Im nächsten Schritt wird dargelegt, welche Akteur:innen diese Debatte tragen und wie sie zur Ausgestaltung des Diskurses beitragen.

9.2. Akteur:innen der bildungspolitischen Debatte über Künstliche Intelligenz

Die Analyse zeigt, dass politische Entscheidungsträger:innen im untersuchten Material mit 63,2 % der codierten Segmente am häufigsten vertreten sind.[31] Wissenschaftliche Akteur:innen (vor allem die wissenschaftlichen Beratungsgremien EFI, TAB und WR) treten in 24,4 % der Segmente auf, während Gewerkschaften (in der vorliegenden Analyse ausschließlich die GEW) lediglich 3,4 % ausmachen. Verbände wie das HFD sind in weiteren 8,5 % der Segmente vertreten. Dieses Übergewicht politischer Entscheidungsträger:innen verweist zum einen auf ihre institutionelle Rolle in Steuerung und Agenda-Setting, ergibt sich jedoch zum anderen auch aus der Zusammensetzung des Datenkorpus, in dem Parlamentsdokumente sowie Dokumente der GWK und KMK 60,8 % der ausgewerteten Materialien ausmachen (siehe Tabelle 5). Damit ergibt sich für die Untersuchungseinheit der vorliegenden Analyse ein Diskursfeld, in dem Regierungs-, Parlaments- und Parteipositionen dominieren, während Expertise und organisierte Interessen eher punktuell akzentuieren.

Eine vertiefte Differenzierung der Positionen politischer Entscheidungsträger:innen zeigt, dass parlamentarische Gremien wie die Enquete-Kommission sowie der Ausschuss für Bildung, Forschung und Technikfolgenabschätzung in 9,4 % der Segmente vertreten sind. Bund-Länder-Zusammenschlüsse wie KMK und GWK erscheinen in weiteren 8,2 % der Segmente. Die Positionierungen von Regierungen, verstanden als explizite Äußerungen im Namen der jeweiligen Bundes- oder Landesregierung, machen 19,4 % der Segmente aus. Den größten Anteil innerhalb der Akteursgruppe politischer Entscheidungsträger:innen bilden parteipolitische Positionierungen mit 26,6 %. Sie umfassen Beiträge, die im Namen von Fraktionen in Plenardebatten oder Ausschusssitzungen beziehungsweise von Politiker:innen in ihrer Rolle als Parteimitglied

[31] Mehrfachnennungen verschiedener Akteur:innen innerhalb eines Segments waren möglich, beispielsweise wenn Anträge im Namen mehrerer Fraktionen gestellt wurden (z. B. LT-NI_541, 2019) oder ein einheitlicher Standpunkt verschiedener Akteur:innen indirekt wiedergegeben wurde (z. B. GEW_1275, 2021, S. 2).

formuliert wurden und damit sowohl regierungstragende als auch oppositionelle Parteien einschließen (zur eingeschränkten Differenzierung von Regierung und Opposition siehe Fußnote 29). Am häufigsten äußerten sich Vertreter:innen der SPD mit 6,7 % und der Union mit 6,3 % in den codierten Segmenten, gefolgt von Bündnis 90/Die Grünen (5,2 %) und FDP (5,1 %). Die Linke (2,2 %) und die AfD (2,0 %) treten deutlich seltener in Erscheinung. Die Verteilung verweist auf eine verstärkte Präsenz regierungstragender Parteien, insbesondere in Anbetracht verschiedener großer Koalitionen auf Bundes- und Landesebene im Untersuchungszeitraum, die es für die Einordnung der qualitativen Ergebnisse zu berücksichtigen gilt. Aufgrund der qualitativen Fokussierung der Untersuchung auf die inhaltliche Rekonstruktion von Rechtfertigungs- und Kritikmustern erfolgt jedoch keine systematische akteursspezifische Analyse im Sinne einer inferenzstatistischen Auswertung, die etwa eine Standardisierung entlang von Redezeiten oder Anteilen an der Gesamtdebatte erfordern würde. Limitierend gilt für die vorliegende Analyse daher, zu berücksichtigen, dass die an dieser Stelle dargestellten Verteilungen nur begrenzte Aussagen über die quantitative Verteilung der Akteursbeiträge an der gesamten bildungspolitischen Debatte zulassen, zumal diese hier relational zu den codierten Segmenten zu verstehen sind. Nichtsdestotrotz können punktuell Differenzen der Positionierungen verschiedener Akteur:innen auch auf qualitativer Ebene hervorgehoben werden, die im Zuge der nachfolgenden Ergebnispräsentation durch die transparente Darstellung der rechtfertigenden und kritisierenden Akteur:innen vertieft wird.

Dabei kann bereits eine Analyse der Überschneidung verschiedener Codierungen Differenzen in der Positionierung verschiedener Akteur:innen sichtbar machen. Ein erster Blick auf Schnittmengen mit der Valenz der geäußerten Rechtfertigungen zeigt etwa, dass die Debatte im untersuchten Material über nahezu alle Akteursgruppen hinweg überwiegend (abwägend) positiv gegenüber dem KI-Einsatz ausfällt (siehe Anhang A7 sowie Abbildung 5). Ausnahmen bilden jedoch Die Linke, deren Positionierungen zu 51,6 % (abwägend) negativ ausfallen, sowie die GEW, die zu 82,9 % kritisch positioniert ist. Die AfD weist hingegen eine auffällige Unschärfe in ihrer Positionierung auf: 64,3 % der zugeordneten Segmente sind (abwägend) positiv, 35,7 % (abwägend) negativ. Diese Ambivalenz deutet weniger auf eine kohärente inhaltliche Linie hin, sondern kann als Indiz einer opportunistischen Rechtfertigungsstrategie interpretiert werden, die je nach Kontext zwischen technologiefreundlichen und alarmistischen Deutungsmustern wechselt, wie nachfolgend vertieft dargestellt wird.

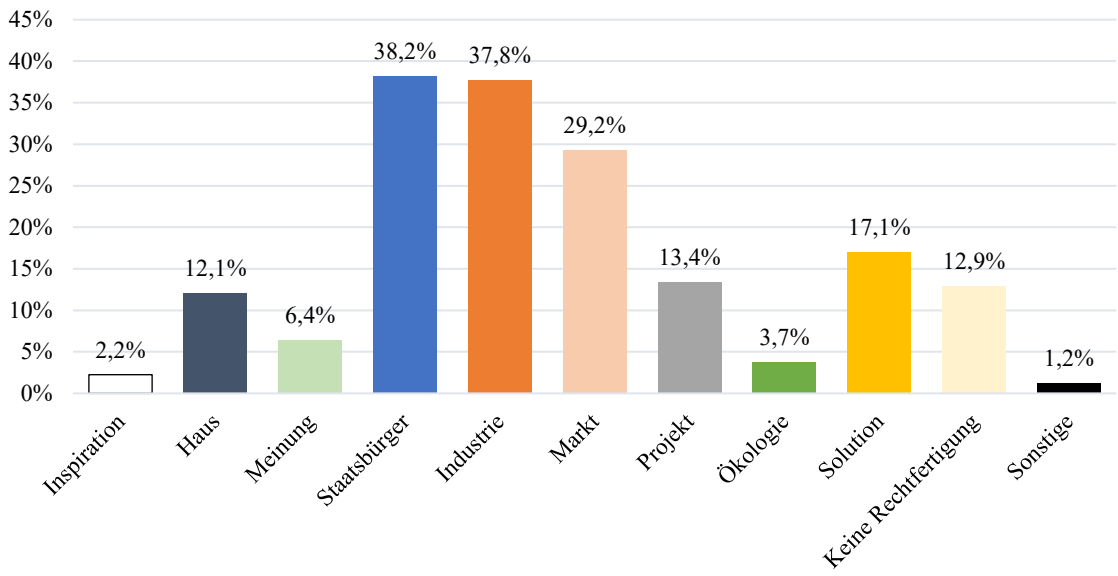
Auf Basis der bisher vorgenommenen Kontextualisierung der bildungspolitischen KI-Debatte, die sowohl die thematischen Bezüge als auch die beteiligten Akteursgruppen erfasst hat, liegt nun die Grundlage vor, um die moralische Fundierung der Debatte vertieft zu analysieren.

9.3. Die moralische Fundierung der Rechtfertigung des Einsatzes Künstlicher Intelligenz

Mit der Analyse der moralischen Fundierung der bildungspolitischen Debatte wendet sich die Arbeit nun der systematischen Beantwortung der Forschungsfragen zu. Im ersten Schritt werden die zur Rechtfertigung und Kritik herangezogenen Cités quantifizierend betrachtet, um zentrale Schwerpunkte der Rechtfertigung des KI-Einsatzes sichtbar zu machen. Im Zuge dessen wurden n = 2.213 Cité-Bezüge in n = 1.067 Segmenten codiert. Methodisch ist daher zu beachten, dass die vorgebrachten Rechtfertigungen nicht durchgehend trennscharf im Sinne eines exklusiven Cité-Bezugs pro Segment erscheinen. Mehrere Cités können in einer Sinneinheit zugleich aufgerufen werden, etwa in Form eines Kompromisses, der auf einer multidimensionalen Rechtfertigung beruht (siehe Abschnitte 9.3.3, 9.4.6 und 9.4.7), oder in offenliegenden Konflikten zwischen unterschiedlichen moralischen Standpunkten (siehe Abschnitt 9.5). Für die Analyse wurden daher Mehrfachcodierungen je Segment zugelassen. Die identifizierten Muster werden im anschließenden qualitativen Teil systematisch rekonstruiert und dabei in Beziehung zu verschiedenen Prüfungsformen, zu den entstehenden Konfliktlinien sowie zu möglichen Kompromissen gesetzt.

Die Häufigkeitsverteilung der Cités innerhalb der einzelnen Segmente (siehe Abbildung 4) zeigt eine deutliche Konzentration auf die Cités des Staatsbürgers (38,2 %) und der Industrie (37,8 %), die jeweils in mehr als einem Drittel aller Segmente als Grundlage der Rechtfertigung und Kritik des KI-Einsatzes herangezogen werden. Ebenfalls häufig tritt die Cité des Marktes auf, die in 29,2 % der Segmente vertreten ist. Solutionistische Rechtfertigungen tauchen in gut jedem sechsten Segment auf (17,1 %) und liegen damit noch vor projektbasierten Rechtfertigungen (13,4 %) sowie Bezügen auf die Cité des Hauses (12,1 %). Bemerkenswert ist der Anteil von Segmenten ohne Rechtfertigung, der immerhin noch 12,9 % der bildungspolitischen KI-Debatte ausmacht. Seltener finden sich die Cité der Meinung (6,4 %) sowie Bezüge auf die Cité der Ökologie (3,7 %) und der Inspiration (2,2 %). Andere Formen der Rechtfertigung (Sonstige) spielen nur eine marginale Rolle (1,2 %). Eine qualitative Prüfung ergab keine Hinweise auf die Emergenz neuer, in der Literatur bislang nicht erfasster Cités. Insgesamt prägen damit staatsbürgerliche, industrielle und marktliche Rechtfertigungsmuster die bildungspolitische KI-Debatte, während ökologische und inspiratorische Bezugnahmen randständig bleiben. Für die weitere Analyse werden ausschließlich jene Cités berücksichtigt, die mindestens fünf Prozent der Segmente betreffen, sodass Bezüge auf die Cité der Inspiration und der Ökologie in der weiteren Darstellung unberücksichtigt bleiben.

Abbildung 4. Relative Häufigkeitsverteilung der Cités auf Segmentebene



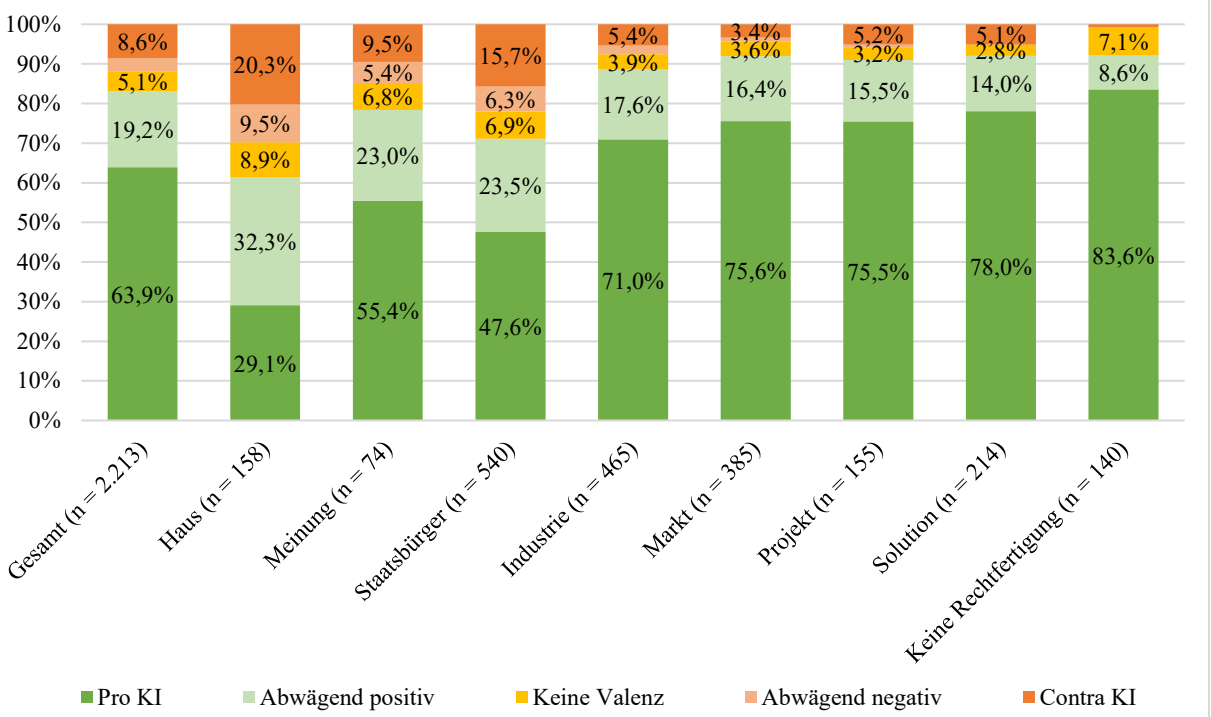
Anmerkung: Analyse auf Segmentebene (n = 1.067). | Mehrfachcodierungen möglich.

Über die relative Häufigkeit der Cité-Bezüge hinaus ermöglicht die Analyse ihrer Valenz sowie der mit ihnen verknüpften Prüfungsformen weitere Erkenntnisse, die das Verständnis der bildungspolitischen KI-Debatte vertiefen. Zunächst werden jedoch Differenzen der moralischen Grundlagen der Rechtfertigungen zwischen den Akteur:innen, insbesondere den Parteien, sichtbar (siehe Anhang A8). Die Cités des Staatsbürgers, der Industrie und des Marktes bilden nicht nur die drei zentralen moralischen Bezugspunkte der Gesamtdebatte, sondern strukturieren ebenfalls die Rechtfertigungen politischer und wissenschaftlicher Akteur:innen sowie des HFDs. Abweichungen zeigen sich insbesondere bei der GEW und bei der Links-Partei, die sich insgesamt in mehr als der Hälfte ihrer Bezugnahmen auf die Cité des Staatsbürgers (GEW: 50,0 %, Die Linke: 38,8 %) oder des Hauses (GEW: 17,7 %, Die Linke: 16,4 %) stützen. Für die AfD kann die Cité des Hauses als dritthäufigster Bezugspunkt (13,7 %) neben der Cité des Staatsbürgers (28,8 %) und des Marktes (27,4 %) nachgewiesen werden. FDP und Union verknüpfen ihre Positionierungen hingegen stärker als andere Akteur:innen mit der Cité der Solution (FDP: 18,5 %, CDU: 17,4 %). Und schließlich lassen sich fehlende Rechtfertigungen besonders häufig in Bezugnahmen von Bundes- und Landesregierungen nachweisen, deren Standpunkte zu 16,6 % ohne weitere Rechtfertigung bleiben.

Ergänzend bestätigt ein differenzierter Blick auf die Valenz der einzelnen Cités zunächst die im vorausgehenden Abschnitt angeschnittene, überwiegend optimistische Rahmung der KI-Debatte. Von den insgesamt n = 2.213 Cité-Bezügen erweisen sich 83,1 % als Befürwortung des KI-Einsatzes oder als abwägend positive Bewertung. Die Analyse verschiebt damit den Fokus von den Segmenten auf die einzelnen moralischen Bezugnahmen und zeigt, dass kritische

Perspektiven, die sich gegen einen KI-Einsatz aussprechen, in der Minderheit bleiben. Diese sind vor allem in der Cité des Staatsbürgers und der Cité des Hauses verortet (siehe Abbildung 5). Abwägende Beiträge ohne klare Valenz markieren eher die Ausnahme. Die somit skizzierte Richtung der bildungspolitischen KI-Debatte kann als ein erstes Indiz dafür betrachtet werden, dass diese überwiegend zustimmend und vergleichsweise konfliktarm geführt wird.

Abbildung 5. Valenz der Rechtfertigungen nach Cités



Anmerkung: Aus Darstellungsgründen wurden nicht alle Datenpunkte beschriftet. Die fehlendenWerte ergeben sich aus dem fehlenden Anteil von Hundert. | Die Stichprobengrößen geben die absoluten Häufigkeiten aller codierten Cités an. Die prozentualen Angaben beziehen sich auf die relative Häufigkeit der Valenz innerhalb dieses Stichprobe.

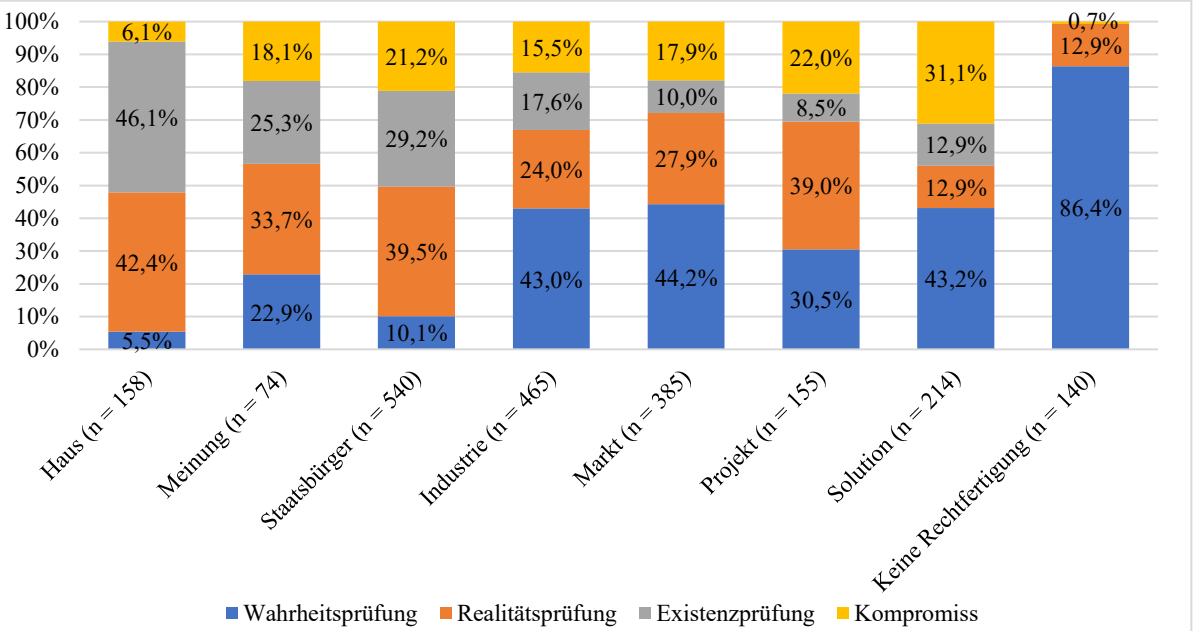
Darauf deutet ebenfalls der hohe Anteil an Wahrheitsprüfungen hin, der eng mit dem positivistischen KI-Bild verknüpft ist (zum Verständnis der Prüfungsformen siehe Tabelle 1). Auf Ebene der Segmentanalyse können in über der Hälfte aller Segmente Wahrheitsprüfungen nachgewiesen werden (51,7 %), gefolgt von Realitätsprüfungen mit 44,4 %. Existenzprüfungen, die Konflikte zwischen unterschiedlichen Cités sichtbar machen, finden sich hingegen nur in knapp einem Fünftel aller Segmente (19,1 %). Kompromisse, die eine gemeinsame Rechtfertigung mehrerer Cités beabsichtigen, treten in 15,9 % der Segmente auf. Da verschiedene Prüfungsformen innerhalb einer Sinneinheit identifiziert wurden, werden Mehrfachcodierungen berücksichtigt. Aufgeschlüsselt nach Cité-Bezügen (siehe Abbildung 6)[32] zeigt sich, dass die Cités der

[32] An dieser Stelle gilt es zu beachten, dass Existenzprüfungen und Kompromisse Mehrfachcodierungen darstellen, da sie immer mehrere Cité-Bezüge gleichzeitig umfassen. Wahrheits- und Realitätsprüfungen markieren dagegen jeweils einen einzelnen Cité-Bezug. Aufgrund dieser strukturellen Differenz wird auf eine aggregierte Darstellung der relativen Häufigkeiten aller Prüfungsformen über sämtliche Cité-Bezüge hinweg verzichtet, da Existenzprüfungen und Kompromisse in einer solchen Gesamtübersicht überproportional erscheinen würden.

Industrie, des Marktes und der Solution besonders häufig im Modus der Wahrheitsprüfung operieren und ihre Rechtfertigungen in mehr als 40 % der Bezüge auf Tatsachenbehauptungen, Wirksamkeits- oder Nutzenargumente stützen. Realitätsprüfungen, die auf eine Diskrepanz zwischen den normativen Erwartungen einer Cité und den tatsächlichen Bedingungen der KI-Implementierung verweisen und damit politische Anpassungsbedarfe markieren, treten besonders häufig in Bezug auf die Cité des Hauses, des Staatsbürgers, des Projekts und der Meinung auf. Existenzprüfungen konzentrieren sich hingegen vor allem auf die Cités des Hauses und des Staatsbürgers und verweisen damit auf normative Zielkonflikte mit anderen Cités. Auffällig erscheint zudem der hohe Anteil an Kompromissen in Verbindung mit solutionistischen Rechtfertigungen (31,12 %), was darauf hindeutet, dass technikbezogene Zukunftsversprechen häufig mit anderen moralischen Ordnungen verknüpft werden. Diese quantitativen Befunde markieren erste Hinweise auf zentrale Mechanismen der Rechtfertigung und Kritik des KI-Einsatzes. Um den dahinterliegenden Bedeutungsstrukturen und der Komplexität der KI-Debatte gerecht zu werden, werden die erkennbaren Muster im folgenden qualitativen Teil vertiefend rekonstruiert.

Im Folgenden werden die aufgeworfenen Forschungsfragen systematisch beantwortet. Zunächst wird die moralische Grundlage der Rechtfertigung des KI-Einsatzes entlang der verschiedenen Cités rekonstruiert (FF1). Den Beginn bilden die zentralen Bezugnahmen auf die Cités der Industrie, des Marktes und des Staatsbürgers (Abschnitte 9.3.1 – 9.3.3). Dabei werden sowohl Wahrheits- und Realitätsprüfungen der einzelnen Cités dargelegt als auch die zwischen diesen Cités erkennbaren Kompromisse dargestellt. Diese entstehen, wenn unterschiedliche moralische Ordnungen gemeinsam zur Legitimation herangezogen werden. Darauffolgend werden die Rechtfertigungen der projektbasierten Cité sowie der Cité der Meinung dargestellt (Abschnitte 9.3.4 und 9.3.5). Anschließend wird untersucht, in welcher Form sich solutionistische Rechtfertigungen in der politischen Debatte über die KI-Einführung manifestieren (FF2; Abschnitt 9.4) und inwiefern ausbleibende Rechtfertigungen als Ausdruck eines deterministischen Optimismus gedeutet werden können (Abschnitt 9.4.5). Zudem werden Kompromissmuster zwischen der Cité der Solution und den Cités der Industrie sowie des Staatsbürgers ausgewiesen (Abschnitte 9.4.6 und 9.4.7). Abschließend wird die moralische Grundlage der Kritik der KI-Einführung analysiert und die daraus entstehenden Konfliktlinien zwischen den Cités im Sinne der Existenzprüfungen herausgearbeitet (FF3; Abschnitt 9.5). In Anbetracht der zuvor dargestellten Verteilung der Cité-Bezüge, ihrer Valenz und Prüfungsformen lässt sich die Kritik des KI-Einsatzes vor allem den Cités des Hauses und des Staatsbürgers zuordnen, weshalb sich die anschließende Ergebnisdarstellung auf diese beiden moralischen Grundlagen konzentriert.

Abbildung 6. Relative Häufigkeit der Prüfungsformen nach Cités



Anmerkung: Die Stichprobengrößen geben die absoluten Häufigkeiten aller codierten Cités an. Die prozentualen Angaben beziehen sich auf die relative Häufigkeit der Prüfungsform innerhalb dieser Stichprobe. | Existenzprüfungen und Kompromisse berücksichtigen mehrfache Cité-Bezüge und werden jeweils für die beteiligten Cités ausgewiesen.

9.3.1. Das industrielle Streben nach Effizienz, Planbarkeit und Innovation

Die Cité der Industrie stellt nach der staatsbürgerlichen Cité den zweithäufigsten moralischen Bezugspunkt der bildungspolitischen KI-Debatte dar. Im Unterschied zur staatsbürgerlichen Perspektive weist sie jedoch mit 88,60 % (abwägend) positiver Bezugnahmen eine ausgesprochen technikoptimistische Grundhaltung auf. Die Cité der Industrie fungiert damit als primäre Grundlage der KI-Rechtfertigung. Charakteristisch ist der hohe Anteil an Wahrheitsprüfungen (siehe Abbildung 6), in denen der Einsatz von KI in einer mitunter deterministischen Technikdeutung durch nahezu als selbstverständlich einhergehende (1) Effizienzgewinne, (2) wachsende Planungssicherheit und (3) technische Modernisierung im Sinne eines Innovationsstrebens gerechtfertigt wird. Ergänzend lassen sich jedoch ebenso Realitätsprüfungen aufdecken, wenn Akteur:innen konkrete Indikatoren und standardisierte Verfahren zur Messung und Evaluation der behaupteten Verbesserungen einfordern.

9.3.1.1. Effizientere und optimierte Bildung durch den Einsatz Künstlicher Intelligenz

Mit nahezu beharrlicher Konstanz wird die Effizienzsteigerung durch die Automatisierung von Prozessen in Wahrheitsprüfungen als zentrales Rechtfertigungsmoment aus Sicht der Cité der Industrie hervorgehoben. Dabei soll der Einsatz von KI Vorteile für alle Bereiche des Bildungswesens generieren, indem etwa Verwaltungs-, Lehr- und Lernprozesse beschleunigt, repetitive Tätigkeiten reduziert und Ressourcen effizienter verwaltet werden (BT_113b, 2020, S. 198; BT_244b, 2020, S. 1; LT-NI_547, 2019, S. 2; LT-SH_670, 2019, S. 42; TAB_1335, 2020, S.

69). Dazu gehören die im Ausschuss für Bildung, Forschung und Technikfolgenabschätzung geforderte Beschleunigung von BAföG-Verfahren durch den KI-Einsatz (BT_455c, 2022, S. 29), die Professionalisierung des Forschungsinformationsmanagements (GWK_1061, 2022, S. 227), die Effizienzsteigerung von Prüfungsverfahren (HFD_1187, 2015, S. 17) oder die Entlastung der Lehrenden von einem breiten Aufgabenspektrum, um individuellere Betreuung zu ermöglichen (BT_288a, 2020, S. 3). Schließlich ließen sich durch den KI-Einsatz auch Effizienzgewinne in Lernprozessen erzielen (KMK_1076, 2021, S. 14; TAB_1335, 2020, S. 140), „indem redundante Inhalte vermieden werden" (LT-SN_894, 2021, S. 2). So betont Abgeordnete Kemmer von der Unions-Fraktion: „Wir sind überzeugt davon, dass beim KI-unterstützten Lernen große Chancen liegen. Richtig angewandt, kann mit KI die Motivation von Schülern gesteigert werden, es lassen sich schnellere, aber auch nachhaltigere Lernerfolge erzielen" (BT_244c, 2019, S. 14625). Schließlich solle das effizientere Lernen einen höheren Studienerfolg bei konsequenterweise niedrigeren Abbruchquoten bewirken (HFD_1206, 2021, S. 1). In diesen Rechtfertigungen kommen ein optimistischer Determinismus sowie eine Outputorientierung zum Ausdruck, die in neoliberaler Tradition daran anknüpfen, messbare Lernerfolge und quantitative Leistungsverbesserungen durch den Einsatz von KI erzielen zu wollen.

Neben quantitativen Verbesserungen wird der KI-Einsatz ebenfalls mit qualitativen Qualitätsgewinnen der Bildung verknüpft (Brandenburg/FDP, BT_199c, 2019, S. 13133; EFI_1082, 2019, S. 94), indem etwa KI-Anwendungen als „Werkzeuge zur Verbesserung der Schülerleistungen" (TAB_1335, 2020, S. 144) eingesetzt werden sollen oder dazu dienen, „die Lehre in Qualität und Didaktik weiterzuentwickeln und Ressourcen effizienter einzusetzen" (EFI_1082, 2019, S. 103). Dabei wird im bildungspolitischen Diskurs ein klares Optimierungsbestreben formuliert, das auf eine „Verbesserung der Qualität, Leistungsfähigkeit und Wirksamkeit der Hochschulbildung durch den Einsatz von KI" (GWK_1052, 2020, S. 1) abzielt. Neben konkreten KI-Anwendungen zeigt die vorliegende Analyse jedoch, dass das Verbesserungsstreben häufig an diffuse Vorstellungen geknüpft ist, wenn beispielsweise „[d]ie Identifikation von geeigneten Anwendungsszenarien, die eine ausreichende Verbesserung von Effizienz, Qualität oder Sicherheit für Verwaltungsleistungen" (BT_162b, 2019, S. 3) erst noch diskutiert wird. Damit verbleiben rechtfertigende Effizienzversprechen häufig auf abstrakter Ebene.

Zugleich äußert sich die industrielle Rechtfertigung jedoch nicht nur in Wahrheitsprüfungen, sondern führt auch zu Realitätsprüfungen, wenn etwa die Frage in den Vordergrund rückt, anhand welcher Indikatoren und Kennziffern sich die angenommenen Verbesserungen messen und überprüfen lassen (Brünler/Die Linke, LT-SN_904, 2021, S. 2788). Trotz des mitunter diffusen Kontextes fordert eine Vielzahl von Akteur:innen klare Kennzahlen, standardisierte

Evaluationen und stabile Rahmenbedingungen, die den Erfolg der vorgesehenen Bildungsmaßnahmen bestimmen und zur Qualitätssicherung und -steigerung beitragen sollen (BT_133b, 2020, S. 206; BT_244b, 2020, S. 1 f.; GWK_1048, 2019, S. 99; HFD_1189, 2016, S. 1; HFD_1191, 2016, S. 22, S. 39; TAB_1335, 2020, S. 25). Damit verschiebt sich die industrielle Rechtfertigung an dieser Stelle zum einen von pauschalen Effizienzannahmen hin zu einer stärker evidenzbasierten Perspektive. Zum anderen wird darin jedoch ebenso eine Rationalisierung von Bildungsprozessen sichtbar, die sich auf den messbaren Teil von Bildung beschränkt.

9.3.1.2. Planungssicherheit und Evaluation durch die Standardisierung künstlicher Intelligenz

Im Anschluss daran rückt dabei in neoliberalistischer Manier die mit dem KI-Einsatz verbundene Standardisierung und Metrisierung in den Fokus. Der Einsatz von KI wird aus Sicht der Cité der Industrie durch steigende Planungssicherheit und Evaluationsmöglichkeiten gerechtfertigt (BT_288a, 2020, S. 3). In dieser Rechtfertigung äußert sich eine zeitliche Perspektive, die darauf abzielt, auf Basis messbarer Leistungsindikatoren sowohl Prognosen zu erstellen und vorausschauend zu antizipieren als auch rückwirkend (Wirksamkeits-)Kontrollen vorzunehmen (BT_157b, 2019, S. 6 f.; BT_288b, 2021, S. 2 f.; GWK_1044, 2018, S. 273; GWK_1056, 2021, S. 151; TAB_1333, 2021, S. 2 f.; TAB_1335, 2020, S. 67 ff.). So sollen etwa „Learning Analytics den Erfolg und Fortschritt des Kurses begleiten“ (BT_288a, 2020, S. 3) oder KI-Anwendungen Prognosen zukünftiger Studierendenzahlen ermöglichen und Personal- beziehungsweise Lehrkapazitäten planen (BT_113b, 2020, S. 325). Ebenso wird betont, dass digitale Tools „Planungssicherheit für die Förderung überjähriger Projekte“ (LT-SH_670, 2019, S. 47) schaffen und damit eine zielgerichtete Steuerung von Investitionen sichern. Diese Rechtfertigungen verweisen damit auf ein deutlich erkennbares Steuerungsinteresse, das KI primär als Instrument der administrativen Kontrolle rahmt.

9.3.1.3. Künstliche Intelligenz als Treiber eines innovativen Bildungssystems und Forderungen nach Investitionen und Fachkräfteentwicklung

Schließlich wird KI unter Bezug auf die Cité der Industrie als Gegenstand der Innovationsförderung im Zuge einer technischen Modernisierung von Lehr- und Lernverfahren sowie der Verwaltungsabläufe gerechtfertigt. Traditionelle, oft veraltete Unterrichts- und Verwaltungsprozesse sollen durch den gezielten Einsatz neuer Technologien ersetzt werden (GWK_1040, 2016, S. 85; GWK_1048, 2019, S. 100; HFD_1214, 2021, S. 3; LT-HE_765, 2020, S. 5; LT-NI_798, 2021, S. 45; Albrecht/Die Grünen, LT-SH_970, 2022, S. 10703) und so zu einer „zeitgemäße[n] Bildung“ (BT_251c, 2020, S. 30) beitragen und zugleich Einsparungen ermöglichen (LT-SH_670, 2019, S. 27; LT-SH_897, 2021, S. 21; TAB_1335, 2020, S. 69). So wird der

technologische Fortschritt als zentrales Mittel angesehen, um den Bildungssektor fit für die Zukunft zu machen (HFD_1187, 2015, S. 47; HFD_1206, 2021, S. 1). Im Kontext der Ausbildung von Fachkräften betont Bildungsministerin Karliczek (CDU) dabei die Notwendigkeit des Einsatzes innovativer Tools, um den Grundstein für die Ausbildung späterer KI-Expert:innen zu legen:

> Es geht um Innovationen *durch* die Ausbildung, aber ebenso geht es um Innovationen *in* der Ausbildung. Die Ausbildung selbst muss neue Technologien wie die Digitalisierung oder auch die künstliche Intelligenz nutzen. Nur so bleiben wir als Innovationsland weiter an der Spitze. (BT_315c, 2020, S. 33; Herv. i. O.)

Vor diesem Hintergrund werden neben der Entwicklung neuer Anwendungen ebenso Investitionen in Infrastrukturen und die KI-Forschung als strategische Maßnahmen im Rahmen von Realitätsprüfungen gerechtfertigt, die als notwendige Bedingung betrachtet werden, um bedeutende Innovationssprünge zu erzielen und den technologischen Wandel an Bildungseinrichtungen voranzutreiben (BT_113b, 2020, S. 333; Schön/Union, BT_113c, 2018, S. 4152; Goldschmidt/Die Grünen, LT-SH_670, 2019, S. 1; LT-NI_798, 2021, S. 80; GWK_1048, 2019, S. 102; WR_1130, 2019, S. 76). Ein „innovationsoffener Ansatz“ (Schön/Union, BT_113d, 2020, S. 23758) wird dabei besonders hervorgehoben, der Kooperationen zwischen Schulen, Hochschulen, außeruniversitärer Forschung sowie Unternehmen bedürfe (BT_113a, 2018, S. 4; Günther/CDU, LT-SH_970, 2022, S. 10720). Im Zuge dessen wird ein „Bedarf an hochqualifiziertem Personal in Wissenschaft und Wirtschaft“ (BT_162b, 2019, S. 11) konstatiert sowie die Notwendigkeit einer „Exzellenzinitiative in der beruflichen Bildung“ (Brandenburg/FDP, BT_244d, 2021, S. 27065). Ein geforderter „Forschergeist“ (FDP, BT_157b, 2019, S. 6) soll durch die Gewinnung ausländischer Fachkräfte und Forschender ebenso wie durch die Ausbildung inländischen Fachpersonals gestützt werden (BT_157a, 2018, S. 4; BT_162b, 2019, S. 20; EFI_1082, 2019, S. 94; LT-HH_930, 2021, S. 1). Innovation wird in diesem Rechtfertigungsmuster vor allem als technologische Modernisierung der Bildung verstanden, deren konkrete Ausgestaltung mitunter vage bleibt. Insgesamt rückt aus Sicht der Cité der Industrie eine stark outputorientierte Perspektive in den Vordergrund, die den KI-Einsatz primär über quantifizierbare Leistungssteigerungen und technologischen Fortschritt rechtfertigt.

9.3.2. Die Zwänge des Wettbewerbs und die Chancen der Individualisierung aus Sicht der Cité des Marktes

Wenngleich die Rechtfertigungen aus Sicht der Cité des Marktes ebenfalls eine ökonomische Ausrichtung erkennen lassen, unterscheiden sie sich von den industriell geprägten Rechtfertigungen. In der Marktlogik rückt KI dabei als (1) Instrument zur Sicherung der

Wettbewerbsfähigkeit, (2) das Versprechen damit verbundener Wohlstandsgewinne sowie (3) wachsende Individualisierung von Bildung zur Herausbildung spezifischer Kompetenzen und persönlicher Vorteile auf dem Arbeitsmarkt in den Vordergrund. Diese Rechtfertigungen manifestieren sich erneut überwiegend in Wahrheitsprüfungen. Ergänzend treten Realitätsprüfungen auf, wenn die Sicherstellung einer ausreichenden Fachkräfteausbildung, internationale Vernetzung und gezielte Investitionen als Bedingungen zur praktischen Wahrung der Wettbewerbsfähigkeit adressiert werden.

9.3.2.1. Künstliche Intelligenz als Garant der Sicherung der Wettbewerbsfähigkeit von Bund, Ländern und Hochschulen

Die marktliche Rechtfertigung des KI-Einsatzes wird maßgeblich durch die Vorstellung geprägt, dass KI einen zentralen Wettbewerbsvorteil für Bund, Länder, einzelne Bildungseinrichtungen und auch Individuen darstellt. KI erscheint als strategische Ressource, deren Nutzung nicht nur die Leistungsfähigkeit des Bildungswesens verbessern, sondern zugleich die Position Deutschlands im globalen Wettbewerb stärken soll (Schipanski/Union, BT_113c, 2018, S. 4258; Steier/Union, BT_157c, 2018, S. 8097; Karlizcek/Union, BT_157d, 2019, S. 9472; GWK_1060, 2022, S. 246; LT-SH_592, 2019, S. 5; LT-BW_877, 2021, S. 33; Gerber/Die Grünen, LT-SN_904, 2021, S. 2788 f.; Knuth/Die Grünen, LT-SH_970, 2022, S. 10723). Biadacz, Abgeordneter der Unions-Fraktion, betont etwa:

> KI ist *das* Schlüsselthema, damit Deutschland und Europa (…) weltweit führende Innovationsstandorte bleiben (…). Wir müssen uns aber in einem harten internationalen Wettbewerb behaupten, damit wir auf dem Gebiet der künstlichen Intelligenz die Zukunft gestalten können. (BT_113d, 2020, S. 23768; Herv. i. O.)

Damit wird KI als Zukunftstechnologie inszeniert, deren Einsatz unverzichtbar sei, um wirtschaftliche Stärke und Bildungsqualität zu sichern. Schleswig-Holsteins Ministerpräsident Günther (CDU) formuliert dies als alternativlos: „Wollen wir auch in Zukunft eine wettbewerbsfähige Wirtschaft; wollen wir künftig zeitgemäße Bildung und exzellente Forschung; (…) dann führt kein Weg an künstlicher Intelligenz vorbei“ (LT-SH_970, 2022, S. 10719), denn KI sei „der Schlüssel für unsere Wettbewerbsfähigkeit“ (ebd., S. 10721).

Gleichzeitig legt die Analyse Differenzen zwischen Regierung und Opposition hinsichtlich der Einschätzung der deutschen Wettbewerbsposition offen. Während etwa die Grünen warnen, Deutschland drohe, „den internationalen Anschluss (…) zu verlieren“ (BT_157a, 2018, S. 2), und die FDP in Bezug auf die KI-Nutzung in der Bildung kritisiert, Deutschland sei „nicht Spitze, sondern ganz weit hinten“ (BT_288b, 2021, S. 13), weist die schwarz-rote

Bundesregierung diese Diagnose zurück und verweist auf bestehende internationale Leistungsfähigkeit, insbesondere durch das DFKI (Schipanski/Union, BT_113c, 2018, S. 4158; Schüle/SPD, BT_157d, 2019, S. 947; BT_288b, 2021, S. 15). Diese Kontroversen verdeutlichen, dass Wettbewerbsrhetorik der Cité des Marktes zwar als gemeinsame moralische Grundlage akzeptiert wird, jedoch unterschiedliche Deutungen darüber bestehen, ob der Status quo als Erfolgsgeschichte oder als Defizit betrachtet wird.

Neben nationalen Ambitionen treten auch innerdeutsche Konkurrenzbeziehungen deutlich hervor. Verschiedene Bundesländer betonen das Ziel, regionale Standortvorteile zu sichern oder auszubauen (Kemmerich/FDP, LT-TH_736, 2020, S. 1525; Althusmann/CDU, LT-NI_776, 2020, S. 8600; LT-NW_895, 2021, S. 144; Wendt/AfD, LT-SN_904, 2021, S. 2791; LT-SH_934, 2021, S. 24; Plambeck/CDU, LT-SH_970, 2022, S. 10721; Sinemus/CDU, LT-HE_1009, 2022, S. 8448). Investitionen in regionale Kompetenzzentren dienen daher nicht nur der Innovationsförderung, sondern explizit der Profilierung im innerdeutschen Wettbewerb (LT-RP_522, 2019, S. 3; LT-BW_793, 2020, S. 12; Dunckel/SPD, LT-SH_830, 2021, S. 8745; Fritzsche/Union, LT-SN_904, 2021, S. 2790). So solle sich etwa „in der Region Stuttgart-Tübingen im Bereich Künstliche Intelligenz ein einmaliges Forschungsökosystem von Weltrang entwickeln“ (GWK_1056, 2021, S. 495).

Auch für Hochschulen wird die KI als Chance zur eigenen Positionierung dargestellt (TAB_1335, 2020, S. 68). So wird die Anwendung von KI bereits als Wettbewerbsvorteil im Hochschulsektor markiert: „Für die Hochschulen bedeutet die Anwendung dieses Prinzips eine größere Zufriedenheit der Studierenden, was wiederum ein Wettbewerbsvorteil (…) sein kann“ (HFD_1187, 2015, S. 53). Hochschulen wie das Karlsruher Institut für Technologie oder die Hochschule RheinMain investieren daher gezielt in KI-Infrastrukturen, um ihr Profil zu schärfen (LT-HE_765, 2020, S. 43; LT-BW_793, 2020, S. 9). Dies spiegelt sich auch in der Zielsetzung der Bund-Länder-Initiative wider, „die Hochschulen zu Vorreitern bei der Anwendung von Künstlicher Intelligenz zu machen“ (GWK_1058, 2021, S. 1). Damit zeigt sich, dass marktliche Rechtfertigungen den KI-Einsatz vor allem über das Versprechen ökonomischer Zukunftsfähigkeit und standortpolitischer Vorteile rechtfertigen. Wettbewerb fungiert dabei als zentraler Maßstab, an dem sowohl nationale als auch regionale und institutionelle Interessen ausgerichtet werden.

9.3.2.2. Internationale Vernetzung und Fachkräfteausbildung als Voraussetzung der Wohlstandsgewinne durch den Einsatz Künstlicher Intelligenz

Eng mit der Wettbewerbsfähigkeit verknüpft ist die Rechtfertigung des KI-Einsatzes durch volkswirtschaftliche Wohlstandsgewinne. KI erscheint nicht nur als strategischer Wettbewerbsvorteil für Bund und Länder, sondern auch als Treiber gesamtwirtschaftlicher Wertschöpfung und individueller Erwerbschancen, etwa durch die Vermittlung von KI-Kompetenzen (Kluckert/FDP, BT_113d, 2020, S. 23765; LT-NI_660, 2019, S. 2; Günther/CDU, LT-SH_830, 2021, S. 8742; LT-SH_934, 2021, S. 7). AfD-Abgeordneter Peschel untermauert die wirtschaftliche Bedeutung von KI dabei durch eine Prognose der erwarteten Gewinne: „Das deutsche Bruttoinlandsprodukt dürfte aufgrund KI-basierter Lösungen bis zum Jahr 2030 um mehr als 11 % steigen. In Euro ausgedrückt sind das ungefähr 430 Milliarden Euro“ (LT-SN_904, 2021, S. 2787). In entsprechenden Rechtfertigungen werden wirtschaftliche Vorteile im Zuge von Wahrheitsprüfungen als nahezu zwangsläufige Folge des KI-Einsatzes gerahmt, wenngleich mitunter unklar bleibt, wie diese Gewinne konkret erzielt werden sollen.

Gleichzeitig zeigt die Analyse, dass mit Forderungen nach internationaler Kooperation und einer qualifizierten Fachkräftebasis auch Herausforderungen im Sinne von Realitätsprüfungen zur Sicherung dieser Wohlstandsgewinne und der Wettbewerbsfähigkeit verknüpft werden. Dabei wird im Zuge der internationalen Vernetzung insbesondere die europäische Zusammenarbeit hervorgehoben (BT_157b, 2019, S. 7; BT_162b, 2019, S. 18 f.), um so etwa „mit den USA und China als den führenden Nationen auf dem Gebiet der KI-Forschung mithalten zu können“ (EFI_1082, 2019, S. 8), wie das EFI betont. Auch Bildungsinstitutionen sollen von internationalen Kooperationen profitieren (WR_1093, 2016, S. 89). Dabei werden sowohl die Gewinnung und Bindung von Spitzenforscher:innen und Expert:innen aus dem In- und Ausland als auch die Rekrutierung des wissenschaftlichen Nachwuchses betont (BT_157a, 2018, S. 4; BT_157b, 2019, S. 7; GWK_1051, 2020, S. 8; GWK_1060, 2022, S. 345). Darüber hinaus spielt die Fachkräfteentwicklung eine zentrale Rolle in marktlogischen Rechtfertigungen, wobei Bildungsministerin Prien (CDU) betont: „Die Hochschulen haben (…) eine herausragende Bedeutung, indem sie die akademischen Fachkräfte ausbilden und bereitstellen sowie den Einsatz von Künstlicher Intelligenz in der Hochschulbildung fördern“ (LT-SH_787, 2020, S. 3). Fachkräfteentwicklung erscheint somit als Voraussetzung dafür, Wettbewerbsvorteile und Wohlstandsgewinne langfristig sichern zu können. Zugleich erfordere die Aufrechterhaltung der Wettbewerbsfähigkeit zwar Investitionen in die KI-Forschung und ihre Anwendung, allerdings wird eine „Gießkannenpolitik“ der KI-Förderung, wie sie teilweise in der unspezifischen industriell geprägten Forderung der innovationsoffenen Förderung zum Ausdruck komme,

abgelehnt (EFI_1082, 2019, S. 8). Stattdessen werden gezielte Förderungen als gewinnbringend erachtet (Wendt/AfD, LT-SN_904, 2021, S. 2791).

9.3.2.3. Individualisierung der Bildung als persönlicher Wettbewerbsvorteil

Neben nationalen und regionalen Wettbewerbsargumenten richtet sich die marktliche Rechtfertigung des KI-Einsatzes auf die individuelle Wettbewerbsfähigkeit. Digitale und KI-bezogene Kompetenzen werden als Voraussetzung beschrieben, „um gute Chancen auf dem Arbeitsmarkt zu haben und um die Gesellschaft, um Demokratie und Wohlstand im 21. Jahrhundert zu erhalten und weiterzuentwickeln“ (SPD/CDU, LT-NI_541, 2019, S. 1). Diese Perspektive rahmt KI als Mittel, das individuelle Bildungslaufbahnen strategisch optimiert.

Vor diesem Hintergrund bildet die Individualisierung des Lernens eine weitere zentrale marktliche Rechtfertigung. Während die Cité der Industrie primär auf eine Steigerung der Bildungsquantität und -qualität im Aggregat zielt, betont die Cité des Marktes die Möglichkeit, Bildungsangebote passgenau auf Einzelne zuzuschneiden. Der Einsatz von KI wird als Chance dargestellt, heterogene Bedarfe präzise zu erfassen, Lernwege flexibel anzupassen und individuelle Förderung zu ermöglichen (BT_288b, 2021, S. 13; HFD_1191, 2016, S. 22; HFD_1194, 2017, S. 1 f.; LT-NI_568, 2019, S. 1; TAB_1335, 2020, S. 67; WR_1177, 2022, S. 38). Ziel sei es, Lernende „weiter zu fördern und zu fordern, ohne den gleichen Anspruch an die gesamte Klasse zu stellen“ (LT-SN_894, 2021, S. 4). Diese Rechtfertigung verweist auf eine markttypische Nachfrageorientierung, bei der Bildungsleistungen zunehmend als individualisierbarer Service erscheinen. KI wird dabei in der bildungspolitischen Debatte als Instrument gerahmt, welches analog zu kundenorientierten Marktangeboten eine flexible Anpassung der Lernumgebung an die persönlichen Fähigkeiten und die Lerngeschwindigkeit jedes Einzelnen biete (HFD_1196, 2018, S. 6; KMK_1076, 2021, S. 5; LT-NW_753, 2020, S. 177; Voigt/CDU, LT-TH_850, 2021, S. 3551). So ermögliche etwa der Einsatz von Virtual Reality „ein an das individuelle Tempo angepasstes Lernen“ (WR_1177, 2022, S. 41). Diese individualisierungsbezogenen Rechtfertigungen des KI-Einsatzes werden im Zuge einer weiteren Flexibilisierung durch orts- und zeitunabhängige Lehr- und Lernszenarien verstärkt (HFD_1187, 2015, S. 47; TAB_1334, 2021, S. 12; TAB_1335, 2020, S. 13). Damit greifen Akteur:innen erneut Marktlogiken auf, die auf bedarfsgerechte, kontinuierlich verfügbare Angebote abzielen. Gerade KI-basierte Tutorensysteme werden dabei hervorgehoben (HFD_1202, 2021, S. 31), wobei etwa „Interaktionen mit dem System, abhängig von den Fähigkeiten, dem Selbstvertrauen und der Geschwindigkeit des Einzelnen, alle 15 - 180 Sekunden stattfinden“ (LT-SN_894, 2021, S. 2).

Diese Befunde verdeutlichen, dass Rechtfertigungen der Cité des Marktes den KI-Einsatz vor allem als Instrument nationaler und regionaler Wettbewerbsstärkung sowie als Service zur Individualisierung des Lernens deuten. Im nächsten Schritt rücken nun die staatsbürgerlichen Rechtfertigungen in den Blick, die eine normative Gegenperspektive eröffnen und stärker auf demokratische Werte, Grundrechte und gesellschaftliche Gleichheit abstellen, zugleich jedoch Kompromisse mit den Cités der Industrie und des Marktes offenbaren.

9.3.3. Staatsbürgerliche Gemeinwohlorientierung als legitimatorische Grundlage des Einsatzes Künstlicher Intelligenz

Die Cité des Staatsbürgers bildet, gemessen an ihrem Gesamtanteil, den zentralen moralischen Bezugspunkt der bildungspolitischen KI-Debatte. Eine differenzierte Betrachtung von Valenz und Prüfungsformen zeigt jedoch, dass sie im Vergleich zu anderen Cités häufiger KI-kritische Positionen markiert (siehe Abschnitt 9.3). Gleichwohl offenbart die Analyse, dass sich auch in diesem Kontext Rechtfertigungen identifizieren lassen, die den KI-Einsatz befürworten. Diese beziehen sich vor allem auf KI als Mittel zur (1) Gemeinwohlförderung, (2) Stärkung von Teilhabe und Mitbestimmung sowie (3) Förderung von Inklusion. Darüber hinaus werden Kompromisse sichtbar, in denen staatsbürgerliche Rechtfertigungen mit marktlichen und industriellen Bezügen verknüpft werden. Die Analyse macht deutlich, dass diese moralischen Ordnungen nicht ausschließlich im Konflikt zueinanderstehen (siehe Abschnitt 9.5.1), sondern in der bildungspolitischen KI-Debatte häufig miteinander verschränkt werden.

9.3.3.1. Künstliche Intelligenz als Instrument der Gemeinwohlförderung auf Basis ethischer Prinzipien

Im Zentrum der Rechtfertigung der Cité des Staatsbürgers steht KI als Instrument, das dem Gemeinwohl dienen solle. Dabei betont etwa die Nichtregierungsorganisation AlgorithmWatch in Bezug auf den Digitalisierungsgesetzesentwurf des Landes Schleswig-Holstein die Rechtfertigungspflicht politischen Handelns: „Staatliche Organe (…) müssen ihre Tätigkeiten stets mit dem Interesse des Gemeinwohls rechtfertigen, damit sie legitim erscheinen“ (LT-SH_929, 2021, S. 23). Entsprechend formuliert die Bundesregierung das Ziel der „Förderung von gemeinwohlorientierten, digitalen Innovationen, die den Menschen in den Mittelpunkt stellen“ (BT_162b, 2019, S. 5). KI-Anwendungen werden somit als gesellschaftliches Nutzeninstrument dargestellt, welches dem Wohlergehen der Menschen dienlich sei. Im Zuge dessen heben verschiedenste Akteursgruppen die Definition ethischer Prinzipien und die Berücksichtigung zentraler Normen und Werte hervor, wie zum Beispiel Diskriminierungsfreiheit, Transparenz, Datenschutz oder das Recht auf Selbstbestimmung und Freiheit (Biadacz/Union, BT_113d, 2020, S. 23769; Jongen/AfD, BT_113d, 2020, S. 23762; Sitte/Die Linke, BT_113d, 2020, S. 23761; Karliczek/Union, BT_157d, 2019, S. 9472; HFD_1201, 2021, S. 5; HFD_1202, 2021,

S. 54; LT-SH_897, 2021, S. 16, 139). Dabei könne auch der KI-Einsatz selbst diese Prinzipien erfüllen. So heißt es etwa, durch den Einsatz von Learning Analytics seien „Bewertungen (…) objektiver zu gestalten, da Lehrende weniger stark beteiligt sind" (HFD_1210, 2021, S. 2), oder dass Algorithmen dazu beitragen, Verwaltungsabläufe „transparenter und bürgerfreundlicher" (LT-SH_934, 2021, S. 6) zu gestalten.

Nichtsdestotrotz bleibt die in der Analyse weitverbreitete und im Rahmen von Wahrheitsprüfungen vorgebrachte Forderung ethischer Prinzipien nicht selten abstrakt. So werden zwar staatsbürgerliche Ansprüche an „Freiheit" oder „Chancengerechtigkeit" (Promny/FDP, LT-HE_1011, 2022, S. 8583) formuliert, die konkrete Ausgestaltung der Erreichung dieser Ideale bleibt jedoch vielfach offen. Ebenso wird „eine Regulierung des Einsatzes von KI im Einklang mit den Prinzipien, Werten und Grundrechten der europäischen Gemeinschaft" (Kuchta/SPD, LT-RP_933, 2021, S. 4) befürwortet, ohne jedoch über den Status einer vagen Beschreibung hinauszugehen. Dies verweist auf ein Rechtfertigungsmuster, das zwischen gemeinwohlorientierten Ambitionen und praktischer Unschärfe oszilliert. Zugleich ist ein zentrales Ergebnis der Analyse, dass die mitunter diffusen Vorstellungen einer an ethischen Prinzipien ausgerichteten KI-Entwicklung in Kompromisse mit industriellen sowie marktorientierten Rechtfertigungen eingebunden werden.

Der staatsbürgerlich-industrielle Kompromiss: Gemeinwohl als innovationsfördernder Rahmen

Im Zentrum des staatsbürgerlich-industriellen Kompromisses steht eine strategische Verflechtung von Gemeinwohlorientierung und industriellen Innovationslogiken, wie sie etwa in der Positionierung Meyers (CDU) zum Ausdruck kommt:

> Ich glaube, es ist wichtig, zu betonen, dass der Fokus darauf liegen muss, dass die künstliche Intelligenz dem Menschen dient. Sie muss entlasten und unterstützen, gleichzeitig aber ethisch und verantwortungsbewusst eingesetzt werden. Das ist die Maxime, die über allem steht. (LT-SN_904, 2021, S. 2786)

Im Zuge dessen stellen etwa Ethik und Datenschutz nicht nur normative Leitlinien dar, sondern werden selbst als Innovationsvorteil markiert. So betont die schleswig-holsteinische Landesregierung, europäische Datenschutzstandards machten Deutschland zu einem „Pionier beim Einsatz von KI" (LT-SH_670, 2019, S. 58). Ebenso werden Forderungen nach einer effektiveren Verwaltung mit dem Anspruch auf Transparenz und Datenschutz flankiert (Knuth/Die Grünen, LT-SH_970, 2022, S. 10700), wobei staatsbürgerliche Ansprüche als Legitimationsgrundlage industrieller Automatisierungsinteressen genutzt werden (FDP, BT_244b, 2020, S. 5 f.). Deutlich wird diese kompromissorientierte Logik auch im Vorschlag der FDP, Einwilligungen zur

Datenverarbeitung im Rahmen von Bildungstechnologien nicht einzeln, sondern über transparente Verfahren zu gestalten, um sowohl die Effizienz als auch die Datensouveränität zu erhöhen (BT_244a, 2019, S. 3). Die Orientierung an ethischen Prinzipien wird im Zuge dessen etwa vom TAB nicht als Hürde, sondern als „innovations- und vertrauensfördernde[r] regulatorische[r] Rahmen" (TAB_1324, 2022, S. 4) bewertet.

Darüber hinaus wird das Innovationsstreben, welches im Zuge der Wahrheitsprüfungen noch ausschließlich technologisch getrieben wurde, im Zuge dieses Kompromisses mit staatsbürgerlichen Werten verknüpft (Mohrs/SPD, BT_113d, 2020, S. 23768; Die Grünen, BT_157b, 2019, S. 7). So verweist etwa die Enquete-Kommission darauf, Innovation sei auf exzellente Grundlagenforschung und den Transfer wissenschaftlicher Erkenntnisse „zum Wohle der Zivilgesellschaft und der Wirtschaft" (BT_113b, 2020, S. 102) angewiesen. Insgesamt verdeutlichen diese Kompromisse, dass die Cité der Industrie um eine zusätzliche staatsbürgerliche Legitimation bemüht ist, um ihre Ansprüche geltend zu machen und gesellschaftliche Akzeptanz für ihre Innovationen zu sichern. In einer öffentlichen Stimmung, die neuen Technologien häufig mit Skepsis begegne, rahmt Meyer (CDU) eine gemeinwohlorientierte KI-Entwicklung nicht nur als regulatorische Auflage, sondern als strategische Voraussetzung für Innovationsfähigkeit. Er betont, man wolle das Potenzial einer exzellenten Forschungslandschaft nutzen, um Wertschöpfung und Arbeitsplätze zu schaffen, zugleich aber wisse man, „dass viele Menschen solchen neuen Technologien und Entwicklungen mit Skepsis begegnen" (LT-SN_904, 2021, S. 2786), weshalb „Datenschutz und Datensicherheit, aber letztlich auch diskriminierungsfreie Algorithmen eine große Rolle spielen" (ebd.) müssten. Nur wenn ethische Standards eingehalten werden, könne Deutschland zu einem attraktiven Standort für KI-Fachkräfte und Start-ups werden. Die Cité der Industrie scheint damit auf die Cité des Staatsbürgers als Grundlage einer nachhaltigen Innovationsdynamik angewiesen zu sein.

Der staatsbürgerlich-marktorientierte Kompromiss: Europäische Werte als Wettbewerbsvorteil

Eine ähnliche Verflechtung zeigt sich zwischen staatsbürgerlichen Ansprüchen und Rechtfertigungen der Cité des Marktes. Staatsbürgerliche Werte werden dabei sowohl auf individueller Ebene als Wettbewerbsvorteil betont, etwa im Sinne eines humanistischen Bildungsideals (siehe vertiefend dazu Abschnitt 9.5.1.2), welches eine bessere Vorbereitung auf den Arbeitsmarkt ermöglichen soll (HFD_1211, 2021, S. 10), als auch auf nationaler Ebene. Die Idee einer menschenzentrierten, an europäischen Werten ausgerichteten KI fungiert hier als strategisches Differenzierungsmerkmal einer deutschen und europäischen KI-Politik und Gestaltung im globalen Wettbewerb (BT_113b, 2020, S. 102; Die Grünen, BT_157b, 2019, S. 2, 7; GWK_1058, 2021, S. 1; HFD_1202, 2021, S. 22). So unterstreicht etwa Bildungsministerin Karliczek

(CDU), dass das Bildungs- und Innovationsland Deutschland „Hand in Hand in eine wettbewerbsfähige Zukunft, eine am Menschen ausgerichtete Zukunft“ (BT_244d, 2021, S. 27056) gehe. Ebenso hebt die Enquete-Kommission hervor, dass eine werteorientierte Forschungspolitik ein „weltweit führendes Forschungsökosystem“ (BT_113b, 2020, S. 105) etablieren könne.

Dabei wird häufig eine doppelte Zielsetzung verfolgt: Einerseits solle Deutschland den „Anschluss an die USA und China erlangen“ (Kamann/AfD, BT_113c, 2018, S. 4151), andererseits wird eine eigenständige, werteorientierte Vorreiterrolle gefordert, etwa durch die Entwicklung transparenter und diskriminierungsfreier Anwendungen, die explizit als Abgrenzung gegenüber China und den USA verstanden wird (Jongen/AfD, BT_113d, 2020, S. 23763; Bayaz/Die Grünen, BT_113d, 2020, S. 23765; BT_157a, 2018, S. 7). Ein stabiler Rechtsrahmen, der Ansprüche an Datenschutz, gesellschaftliche Gerechtigkeit und Grundrechte sichert, wird hierbei als Bedingung für eine zukunftsfeste, konkurrenzfähige KI-Entwicklung betrachtet (BT_157a, 2018, S. 1 f.; AfD, BT_157b, 2019, S. 5; WR_1175, 2022, S. 58 f.). Marktlogiken und staatsbürgerliche Normen werden besonders deutlich verschränkt, wenn argumentiert wird, dass nur die Sicherung eines international konkurrenzfähigen Standorts ermögliche, eigene Werte dauerhaft in die KI-Entwicklung einzubringen (Christmann/Die Grünen, BT_157d, 2019, S. 9481). Entsprechend betont das Ministerium für Wissenschaft, Weiterbildung und Kultur des Landes Rheinland-Pfalz, dass das gegenwärtige „Zeitfenster“ genutzt werden müsse, „um die Weichen für eine konkurrenzfähige Forschung und Entwicklung Künstlicher Intelligenz zu stellen und gleichzeitig für Rahmenbedingungen zu sorgen, die einen Einsatz gewährleisten, der unserem demokratischen, ökologischen und sozialen Wertesystem folgt“ (LT-RP_522, 2019, S. 1). Dabei erinnert diese implizierte Dringlichkeit an das technologische Momentum, in dem die Weichen der KI-Entwicklung als zeitlich begrenzt politisch gestaltbar anerkannt werden, während diese zugleich mit der Sicherung internationaler Wettbewerbsfähigkeit verknüpft wird.

Insgesamt zeigt die Analyse, dass die staatsbürgerliche Rechtfertigung von KI als Instrument der Gemeinwohlförderung KI zwar als öffentliches Gut rahmt, die konkrete Ausgestaltung jedoch erneut vage bleibt und auf abstrakten normativen Leitbildern beruht. Zugleich wird der Gemeinwohlbezug strategisch genutzt, um technische Innovation politisch anschlussfähig zu machen und sie mit industriellen und marktwirtschaftlichen Interessen zu verknüpfen.

9.3.3.2. Teilhabe und Mitbestimmung als staatsbürgerliche Rechtfertigung eines legitimen Einsatzes Künstlicher Intelligenz

Das zweite zentrale Rechtfertigungsnarrativ der Cité des Staatsbürgers umfasst den Anspruch gesellschaftlicher Teilhabe und Mitbestimmung als Voraussetzung für die Legitimität des KI-

Einsatzes. Der gemeinwohlorientierte Ansatz erfordert, dass KI als gesamtgesellschaftliche Aufgabe verstanden wird (BT_162b, 2019, S. 5 f.), an der nicht nur staatliche Institutionen oder zentrale Entwicklungseinrichtungen, sondern alle gesellschaftlichen Akteur:innen beteiligt sind, insbesondere Lernende, Lehrende, Forschende und Verwaltungsmitarbeitende als Betroffene des KI-Einsatzes. In dieser Perspektive erscheint die Tradition einer partizipativen Technikgestaltung (siehe Abschnitt 6.3.3) nicht als Ergänzung, sondern als notwendige Bedingung einer wertebasierten KI-Entwicklung (LT-SH_934, 2021, S. 14). Diese solle, analog zu der Ethik-Leitlinie einer vertrauenswürdigen KI der Europäischen Kommission (2019), zugleich Transparenz ermöglichen sowie Akzeptanz und Vertrauen in KI schaffen (BT_113b, 2020, S. 31, 205, 319; Röspel/SPD, BT_113d, 2020, S. 23762; LT-HE_765, 2020, S. 45 f.; LT-BW_793, 2020, S. 6; Gerber/Die Grünen, LT-SN_904, 2021, S. 2791; Holowaty/FDP, LT-SH_970, 2022, S. 10726). Dabei kann diese Bestrebung auf der einen Seite im Sinne einer „intensive[n] gesellschaftliche[n] Debatte“ (Domscheit-Berg/Die Linke, BT_157c, 2018, S. 8101) als beabsichtigte Politisierung verstanden werden (siehe Abschnitt 6.3.4), die im Sinne einer Wahrheitsprüfung durch die alleinige gesellschaftliche Mitbestimmung die Legitimität der KI sichern solle. Auf der anderen Seite kritisiert die GEW jedoch eine bislang mangelnde Politisierung des KI-Einsatzes im Bildungsbereich (GEW_1338, 2021, S. 28). Auch Die Linke fordert deshalb einen „diskursive[n] Beteiligungsprozess innerhalb der Schule“ (BT_195b, 2019, S. 10), um Fehlentwicklungen und Fehlinvestitionen vorzubeugen. Dies gelte insbesondere vor dem Hintergrund, „dass KI gesellschaftsverändernden Charakter hat; auch deshalb muss Öffentlichkeit möglichst früh beteiligt werden“ (BT_113d, 2020, S. 23760), wie Abgeordnete Sitte (Die Linke) darlegt. An dieser Stelle zeigt sich eine Verschmelzung sozialkonstruktivistischer und deterministischer Deutungsmuster, die auch Schröter (Deutscher Gewerkschaftsbund) hervorhebt: „Wir befinden uns an einem Scheideweg, an dem in Bezug auf Algorithmen wichtige Entscheidungen getroffen werden, die nur schwer rückgängig gemacht werden können“ (GEW_1338, 2021, S. 22). Aufgrund dieser Irreversibilität in Kombination mit der Drohung einer wachsenden gesellschaftlichen Spaltung gelte es als Gewerkschaft, „sich intensiv in technische Fragen einzumischen“ (ebd.). Damit wird die staatsbürgerliche Teilhabe gerade aufgrund des Risikos langfristiger Pfadabhängigkeiten gerechtfertigt, worin sich ein weiteres Beispiel für die Anerkennung des technologischen Momentums zeigt.

Die mit einer gemeinwohlorientierten KI-Entwicklung verknüpfte Forderung gesellschaftlicher Teilhabe äußert sich somit sowohl in Wahrheitsprüfungen als auch in Realitätsprüfungen. Verschiedene Akteur:innen betonen in diesem Zusammenhang die Notwendigkeit des Aufbaus gesellschaftlicher KI-Kompetenzen als Voraussetzung einer gleichberechtigten Mitbestimmung

der KI-Gestaltung sowie zur individuellen beruflichen und alltäglichen Anschlussfähigkeit (Bundesregierung, BT_288b, 2021, S. 12; TAB_1335, 2020, S. 18, 140). Der Wissenserwerb über KI solle befähigen, „an den digitalisierten Lebens- und Arbeitswelten" (GEW_1338, 2021, S. 16) aktiv teilzuhaben und „neue soziale Spaltungen" (ebd.) zu vermeiden. Dies gelte Röspel (SPD) zufolge insbesondere für Personen, die von beruflichen Veränderungen oder Substituierungen betroffen sein werden und die es daher verstärkt zu unterstützen gelte (BT_113d, 2020, S. 23762). Dabei wird der Wissenserwerb über KI auch mit staatsbürgerlichen Idealen wie staatlicher und persönlicher Souveränität gerechtfertigt (Die Grünen, BT_157a, 2018, S. 1). In dieser Perspektive soll „Lernen über und (…) mit KI (…) der gesamten Bevölkerung Mündigkeit und Befähigung für den Umgang mit KI vermitteln" (BT_113b, 2020, S. 350). Gleichzeitig wird Kompetenzaufbau auch als Fähigkeit zur kritischen Reflexion technischer Prozesse betrachtet, etwa zur Vermeidung von Fehlinterpretationen algorithmischer Ergebnisse (WR_1175, 2022, S. 23). Die curriculare Berücksichtigung eines verantwortungsvollen KI-Umgangs solle daher sowohl zur eigenständigen Anwendung als auch zur reflektierten Interpretation algorithmischer Outputs befähigen (HFD_1214, 2021, S. 2). Diese Rechtfertigungen verdeutlichen somit, dass Teilhabe in der staatsbürgerlichen Perspektive nicht nur als demokratische Notwendigkeit verstanden wird, sondern auch als präventive Antwort auf digitale Ungleichheit und als Grundlage einer souveränen, kompetenten Bürgerschaft.

Der Kompromiss staatsbürgerlicher Teilhabe als Voraussetzung industrieller Wertschöpfung

Zugleich wird der Anspruch der sozialen Teilhabe mitunter als Voraussetzung für industrielle Wertschöpfung verstanden, was erneut auf einen staatsbürgerlich-industriellen Kompromiss verweist. Ein Beispiel bildet die Ausbildung von Studierenden im Studiengang „Human Factors", die als Vermittler:innen zwischen Mensch und Technik qualifiziert werden sollen. Ziel sei eine möglichst leistungsfähige Systemgestaltung bei gleichzeitiger Sicherstellung des Wohlergehens aller Beteiligten (HFD_1202, 2021, S. 37). Teilhabe wird so nicht nur als demokratisches Ideal, sondern auch als funktionale Bedingung technologischer Effizienz interpretiert. Diese wechselseitige Bezugnahme wird auch in der pädagogischen Praxisentwicklung sichtbar. Die Einbindung von Lernenden und Lehrenden in technologische Gestaltungsprozesse gilt einerseits als Beitrag zur Optimierung industrieller Entwicklungen, andererseits als Mittel zur demokratischen Legitimierung und Qualitätssicherung (LT-HE_765, 2020, S. 10). Die staatsbürgerlich begründete Kompetenz zur kritischen Reflexion wird hierbei als zentral angesehen, um technologische Innovationen sinnvoll in Lernprozesse einzubetten, ihre Effizienz zu evaluieren und sie didaktisch weiterzuentwickeln (KMK_1065, 2017, S. 49).

Kompetenzvermittlung als Kompromiss staatsbürgerlicher und marktlicher Rechtfertigungen

Die Kompetenzvermittlung wird ebenfalls im Kompromiss zwischen staatsbürgerlichen und marktlichen Rechtfertigungen aufgegriffen. Die Förderung von KI-Kompetenzen wird gleichzeitig als wirtschaftliches Gebot und als Voraussetzung gesellschaftlicher Teilhabe dargestellt. In einer digitalisierten Welt, in der nahezu jede:r mit KI-Systemen in Berührung kommt, werden Kenntnisse über deren Funktionsweise und Gestaltung explizit sowohl als Schlüsselqualifikation für individuelle Wettbewerbsfähigkeit und den Erfolg des Wirtschaftsstandorts als auch für demokratische Teilhabe und Selbstbestimmung benannt (BT_113b, 2020, S. 332; TAB_1335, 2020, S. 17). Angesichts der Sorge, dass digitale Ungleichheit ökonomische Ungleichgewichte verstärken könne, wird gesellschaftliche Teilhabe zudem von Abgeordneter Tatti (Die Linke) als Voraussetzung wirtschaftlicher Stabilität und persönlicher Absicherung dargestellt (BT_251f, 2020, S. 18410). Eine ähnliche Perspektive formuliert der Kultusausschuss des Landtages Niedersachsen, der betont, dass alle Bürger:innen digitale Kompetenzen benötigten, „um gute Chancen auf dem Arbeitsmarkt zu haben und um die Gesellschaft, um Demokratie und Wohlstand im 21. Jahrhundert zu erhalten und weiterzuentwickeln“ (LT-NI_660, 2019, S. 2). Der Anspruch, Bildung an den Erfordernissen eines digitalen Arbeitsmarktes auszurichten, wird damit nicht als Gegensatz zu staatsbürgerlichen Zielen verstanden, sondern als deren Voraussetzung. Die Befähigung zur Teilhabe in einer digitalisierten Gesellschaft bildet so die Schnittstelle, an der sich staatsbürgerliche und marktliche Logiken in einem produktiven Kompromiss verbinden.

9.3.3.3. Inklusion und Chancengleichheit als staatsbürgerliche Rechtfertigung des Einsatzes Künstlicher Intelligenz und als Kompromissgrundlage industrieller und marktorientierter Ansprüche

Die staatsbürgerliche Rechtfertigung des KI-Einsatzes betont schließlich auch dessen Beitrag zu Inklusion und Chancengleichheit. So soll KI Lernprozesse und Lesekompetenzen benachteiliger Schüler:innen verbessern, ebenso wie den Lernerfolg in Integrationskursen erhöhen (BR_243a, 2019, S. 4; BT_288a, 2020, S. 2; LT-SH_934, 2021, S. 30; TAB_1327, 2021, S. 2). Das am Karlsruher Institut für Technologie eingesetzte KI-basierte Echtzeit-Übersetzungs- und Transkriptionstool, wird dabei als Beitrag zur Barrierefreiheit hervorgehoben (BT_113b, 2020, S. 298). Ähnliches gelte für die bereits thematisierten (Abschnit 9.1.1) Avatare, die in hessischen Schulen zum Einsatz kommen und längerfristig erkrankten Schüler:innen die Teilnahme am Unterricht ermöglichen sollen (LT-HE_979, 2022, S. 1). Darüber hinaus fließen auch geschlechterpolitische Aspekte in die Debatte ein, wenn Investitionen in KI-Forschung als Möglichkeit dargestellt werden, Frauen gezielt zu fördern und so geschlechterbezogene Ungleichheiten abzubauen (BT_157a, 2018, S. 4; BT_162b, 2019, S. 11).

Effizienz durch Inklusion: Der staatsbürgerlich-industrielle Kompromiss

Auch diese staatsbürgerliche Rahmung der KI als Inklusionsinstrument wird häufig mit industriellen Rechtfertigungen verknüpft. Anwendungen, die ursprünglich im Sinne von Effizienz und Ressourcenschonung entwickelt wurden, wie *Massive Open Online Courses*, werden zusätzlich als niedrigschwellige Bildungsangebote legitimiert, die Zugangsoffenheit und Chancengleichheit fördern (TAB_1335, 2020, S. 60). Auf diese Weise beabsichtigen verschiedene Akteur:innen der bildungspolitischen KI-Debatte, Effizienz- und Inklusionsansprüche in Einklang zu bringen, indem Automatisierungsprozesse zugleich als Beitrag zur Bildungsgerechtigkeit gerahmt werden (BT_113b, 2020, S. 196; KMK_1065, 2017, S. 32; Friedel/SPD, LT-SN_904, 2021, S. 2790; LT-SH_897, 2021, S. 12; TAB_1333, 2021, S. 2, 6; TAB_1335, 2020, S. 140). Somit zeigt sich ein Kompromiss zwischen industrieller Innovationslogik und staatsbürgerlichen Werten, da Effizienzgewinne moralisch durch Inklusionsziele aufgeladen werden.

Chancengleichheit durch Individualisierung: Der staatsbürgerlich-marktliche Kompromiss

Ein weiterer Zusammenhang ergibt sich erneut im Kompromiss zwischen staatsbürgerlicher und marktorientierter Rechtfertigung. Während staatsbürgerlich argumentierende Akteur:innen die durch den KI-Einsatz vorangetriebene Individualisierung teils kritisch als Treiber sozialer Ungleichheit diskutieren (siehe Abschnitt 9.5.1.1), wird im Rahmen marktlicher Rechtfertigungen das emanzipatorische Potenzial individueller Lernpfade hervorgehoben. KI-gestützte Systeme sollen Lernrückstände gezielt abbauen und damit zur Reduktion von Bildungsbenachteiligungen beitragen (BT_113b, 2020, S. 350; FDP, BT_406, 2021, S. 5; HFD_1191, 2016, S. 34). Die Individualisierung wird damit als Möglichkeit präsentiert, Diversität und Heterogenität besser zu berücksichtigen und die Chancengleichheit zu erhöhen (HFD_1201, 2021, S. 2). Marktliche Zielsetzungen einer individuellen Leistungsoptimierung verbinden sich so mit staatsbürgerlichen Forderungen nach Bildungsgerechtigkeit.

Insgesamt zeigt die Analyse der staatsbürgerlichen Rechtfertigungen des KI-Einsatzes, dass diese von gemeinwohlorientierten und inklusiven Wahrheitsprüfungen geprägt sind. Gleichzeitig treten Realitätsprüfungen hervor, die auf ungelöste Herausforderungen in der praktischen Umsetzung verweisen, etwa hinsichtlich gesellschaftlicher Teilhabe, Kompetenzaufbau oder der konkreten Gestaltung einer menschenzentrierten KI-Entwicklung. Die identifizierten Kompromisse zeigen darüber hinaus, dass staatsbürgerliche Werte eine zentrale Rolle bei der moralischen Stabilisierung industrieller und marktlicher Rechtfertigungen spielen und damit den normativen Rahmen der bildungspolitischen KI-Debatte wesentlich prägen.

9.3.4. Die Vernetzung des Bildungssektors aus Sicht der Cité des Projekts

Die Cité des Projekts bildet mit 13,4 % der codierten Segmente einen weniger häufig herangezogenen moralischen Bezugspunkt, gewinnt jedoch dann an Bedeutung, wenn der KI-Einsatz konkret ausgestaltet oder als Teil umfassender Modernisierungsprozesse diskutiert wird. Charakteristisch ist eine überwiegend positive Deutung von KI, die in Wahrheitsprüfungen als Ausdruck eines notwendigen technologischen Wandels erscheint, dem sich moderne Gesellschaften aktiv stellen müssen. In Realitätsprüfungen rücken hingegen praktische Anforderungen in den Vordergrund, etwa der Aufbau tragfähiger Netzwerke, die Entwicklung neuer Curricula und Standards oder internationale Kooperationen. Zwei zentrale Aspekte strukturieren diese projektbasierte Rechtfertigung: KI wird (1) als notwendiges Modernisierungsprojekt verstanden, das kontinuierliche Veränderungsbereitschaft erfordert, und (2) als Instrument, um Flexibilisierung voranzutreiben und Netzwerke im Bildungsbereich aufzubauen.

9.3.4.1. Künstliche Intelligenz als Projekt moderner Gesellschaften

In der projektbasierten Rechtfertigung werden die Entwicklung und Nutzung von KI als gesamtgesellschaftliches Modernisierungsvorhaben gerahmt. Demnach sollen Politik, Wissenschaft, Wirtschaft und Zivilgesellschaft gemeinsam beraten, wie Bildung und Gesellschaft im technologischen Wandel weiterentwickelt werden können (Steier/Union, BT_244c, 2019, S. 14548; BT_288a, 2020, S. 4; LT-HH_497, 2018, S. 1; LT-SH_670, 2019, S. 14; Waldinger-Thiering/SSW, LT-SH_830, 2021, S. 8748; TAB_1335, 2020, S. 151; WR_1166, 2022, S. 22). Diese Bezugnahmen beruhen auf der Annahme eines grundlegenden Wandels, der das Bildungssystem langfristig prägen werde. So heißt es in einem Antrag der niedersächsischen SPD- und CDU-Fraktionen: „Langfristig wird sich das Bildungssystem durch Digitalisierung grundlegend verändern“ (LT-NI_541, 2019, S. 4). Kultusminister Tonne (SPD) bezeichnet den Transformationsprozess zudem als einen „der größten Umbrüche für Schulen in der jüngeren Vergangenheit“ (LT-NI_664, 2019, S. 5804). Daher gelte es, diesen Prozess positiv zu begleiten (LT-NI_541, 2019, S. 4), indem beispielsweise Curricula und Lerninhalte weiterentwickelt (LT-NW_753, 2020, S. 177) und neue Kompetenzen geschaffen würden, wie etwa Brünler (Die Linke) fordert: „Darum muss Schule ein agiles System werden, das unsere Kinder mit allen Hard- und Soft Skills versorgt, die in dieser Welt einer ständigen Veränderung benötigt werden“ (LT-SN_904, 2021, S. 2788). Diese Äußerungen zeigen, wie die projektbasierte Cité Veränderung nicht als Problem, sondern als normativen Anspruch eines modernen Bildungssystems deutet.

Während diese Modernisierungsforderung teilweise den Innovationsansprüchen der industriellen Cité ähnelt, steht hier weniger technologische Leistungssteigerung als vielmehr die

Anpassungsfähigkeit des Bildungssektors im Vordergrund. So gelte es, „auf wichtige Trends und Entwicklungen (…) einzugehen und (…) Deutschland fit für die Zukunft aufzustellen“ (BT_199b, 2019, S. 19). Flexibilität wird damit zur zentralen Anforderung, die etwa sowohl didaktische Konzepte als auch Bedarfsplanungen und Raumstrukturen betrifft (WR_1166, 2022, S. 22; WR_1174, 2022, S. 92). Der digitale Wandel wird dabei als Vielzahl konkreter Einzelprojekte dargestellt, die schrittweise umgesetzt werden müssen (Fritzsche/CDU, LT-SN_904, 2021, S. 2790). Die Thüringer Strategie für die Digitale Gesellschaft verweist entsprechend auf die Notwendigkeit, „neue Wege für die Bildung an Hochschulen zu finden“ (LT-TH_836, 2021, S. 280).

Zugleich zeigen einzelne Beiträge, dass dieser Wandel durchaus konflikthaft wahrgenommen werden kann. Besonders deutlich wird dies im Bundestag, wenn Sattelberger (FDP) den bisherigen Umgang mit digitaler Transformation kritisiert: „Es geht nicht um ein bisschen mehr Onlineunterricht. (…) Was Sie gar nicht können: Krisen managen, Deutschland fit machen für die Digitalisierung, Transformieren, Lösungen finden für die junge Generation. Dieses Land braucht einen Neustart“ (BT_244d, 2021, S. 27059). Diese Kritik markiert eine projektlogische Realitätsprüfung, die Diskrepanzen zwischen dem erwünschten Modernisierungstempo und den tatsächlichen Fortschritten betont.

Trotz solcher Herausforderungen wird der technologische Wandel insgesamt weitgehend als kontrollierbar dargestellt. Regierungsseitige Rechtfertigungen rahmen ihn häufig im Modus der Wahrheitsprüfung, etwa wenn Tonne (SPD) auf die politische Handlungsfähigkeit verweist, wobei der digitale Transformationsprozess handhabbar, wenngleich andauernd sei (LT-NI_664, 2019, S. 5804). Die Enquete-Kommission unterstreicht diese Normalisierung, indem sie den begleitenden Wertewandel als integralen Bestandteil gesellschaftlicher Entwicklung einordnet: „Der parallel mit dem technologischen Wandel einhergehende Wertewandel ist nicht per se schlecht, der Wertewandel gehört zur Entwicklung von Mensch und Gesellschaft“ (BT_113b, 2020, S. 31). Erneut äußert sich darin ein techniksoziologisches Deutungsmuster, indem Veränderung sowohl unvermeidlich wahrgenommen wird als auch gleichzeitig aktiv angenommen und gestaltet werden soll, wodurch der KI-Einsatz als Teil eines kontinuierlichen gesellschaftlichen Projekts legitimiert wird.

9.3.4.2. Netzwerkbildung als Flexibilisierungsanforderung

Darüber hinaus wird der KI-Einsatz aus Sicht der Cité des Projekts durch die Auflösung starrer Strukturen und die Förderung von Autonomie, Kreativität und Flexibilität gerechtfertigt, sodass auch Freiraum zum Experimentieren eingeräumt werden könne. Entsprechend hebt Korkmaz-

Emre (SPD) hervor, digitale Selbstbestimmtheit bedeute, „dass wir es den vielen klugen Köpfen da draußen auch ermöglichen, Dinge einfach mal auszuprobieren, vielleicht zu scheitern, vielleicht Erfolge zu erzielen“ (BT_251f, 2020, S. 18417). Diese Betonung des Experimentellen markiert ein projektlogisches Motiv, indem Wandel nicht nur akzeptiert, sondern aktiv erprobt und gestaltet werden soll.

Im Zentrum einer nach Flexibilisierung strebenden projektbasierten Rechtfertigung steht der Aufbau vielfältiger Netzwerke auf Ebene von Nationalstaaten, Politik, Bildungsinstitutionen, Forschenden, aber auch Lehrenden und Lernenden. KI soll dazu beitragen, neuen Anforderungen an „flexibles und vernetztes Arbeiten“ (TAB_1335, 2020, S. 195) gerecht zu werden und so kollaborative Prozesse und die Bildung von Netzwerken zwischen Lernenden zu ermöglichen (HFD_1191, 2016, S. 26; HFD_1196, 2018, S. 2). Pilotprojekte dienen in dieser Perspektive als Experimentierfelder, die einen weitläufigeren KI-Einsatz vorbereiten können (HFD_1196, 2018, S. 6; Tonne/SPD, LT-NI_664, 2019, S. 5805, TAB_1335, 2020, S. 234). Projektlogisch wird der Bildungssektor so als Feld fortlaufender Erprobungen gerahmt, in dem Best Practice durch sukzessive Iteration entsteht.

Die Netzwerkorientierung bezieht sich jedoch nicht nur auf Lernende, sondern auch auf Bildungs- und Forschungseinrichtungen. So sollen Schulen und Ausbildungsstätten durch digitale Infrastrukturen stärker verbunden werden, etwa durch ortsunabhängiges Lernen (LT-NI_735, 2020, S. 10), während die Vernetzung von Hochschulen das Teilen von Forschungsergebnissen und Nutzungserfahrungen ermöglichen soll (LT-RP_933, 2021, S. 4; Albrecht/Die Grünen, LT-SH_970, 2022, S. 10703). Im Zuge dessen werden von politischen Entscheidungsträger:innen vielfach internationale und interdisziplinäre Forschungskooperationen angeregt (BT_157a, 2018, S. 4; Christmann/Die Grünen, BT_157d, 2019, S. 9480; GWK_1047, 2019, S. 396; GWK_1051, 2020, S. 57; LT-RP_522, 2019, S. 3; LT-NW_753, 2020, S. 31; LT-HE_765, 2020, S. 42 f.; Waldinger-Thiering/SSW, LT-SH_830, 2021, S. 8748; LT-NW_895, 2021, S. 170; WR_1177, 2022, S. 94).

Die aktive Vernetzung wird dabei im Zuge der Cité des Projekts schließlich auch als politische Aufgabe verstanden, die eine Herausforderung im Sinne einer Realitätsprüfung darstellt. So erfordere die technologische Weiterentwicklung des Bildungssystems eine stärkere Vernetzung von Bund, Ländern und Kommunen (Schipanski/Union, BT_244d, 2021, S. 27067). Dabei müsse „der kooperative Föderalismus (…) weiterentwickelt“ (Rupprecht/Union, BT_244d, 2021, S. 27062) und gemeinsame Infrastrukturen für die KI-Nutzung im Bildungswesen geschaffen werden (Tonne/SPD, LT-NI_664, 2019, S. 5805; LT-SH_897, 2021, S. 116). Darüber

hinaus werden auch internationale Kooperationen gefordert, insbesondere zur Schaffung europäischer Rahmenbedingungen des KI-Einsatzes (BT_113a, 2018, S. 3; Bayaz/Die Grünen, BT_113d, 2020, S. 23766; Sauer/Union, BT_113d, 2020, S. 23764; BT_157a, 2018, S. 2; HFD_1214, 2021, S. 2; TAB_1335, 2020, S. 22). Außerdem könne man durch internationale Zusammenarbeit auch von bestehender Expertise und von Best-Practice-Beispielen, wie dem kostenlosen finnischen KI-Lernprogramm, profitieren (BT_288a, 2020, S. 4). Dabei liegt ein zentraler Anspruch der Cité des Projekts auch in der Arbeitsteiligkeit der Netzwerke begründet, von der nicht nur die Forschung (GWK_1044, 2018, S. 289), sondern ebenso die KI-Regulierung profitieren könne, so Karliczek (CDU):

> Unsere Stärke liegt in der Arbeitsteiligkeit. (…) Die vor uns liegende technologische Entwicklung ist so bedeutend, dass wir mindestens auf europäischer Ebene, wenn nicht sogar in der gesamten demokratischen Welt gemeinsam unsere Werte als Leitplanken des technologischen Fortschritts verankern müssen. (BT_157d, 2019, S. 9472)

Die KI-Entwicklung und ihr Einsatz werden somit aus Sicht der Cité des Projektes nicht nur als Motor eines unumgänglichen Transformationsprozesses gerechtfertigt, sondern ebenso als aktives Gestaltungsprojekt einer arbeitsteiligen Gesellschaft. Dabei thematisieren Realitätsprüfungen vor allem die organisatorischen Voraussetzungen für diese Modernisierung und finden die Lösung in der Bildung vielfältiger Netzwerke. Im Vergleich zu anderen Cités liegt die Besonderheit der projektbasierten Rechtfertigung somit darin, technologische Entwicklung weder primär als Effizienzproblem noch als Wettbewerbsfrage zu behandeln, sondern einer Logik des Temporären und Situativen zu folgen, wobei KI-Entwicklung als Abfolge von Projekten verstanden wird, die spezifische Probleme kontextbezogen und netzwerkbasiert bearbeiten.

9.3.5. Künstliche Intelligenz als Imagegewinn und das Akzeptanzproblem der Cité der Meinung

Die Cité der Meinung bewertet KI-Anwendungen danach, ob sie das Ansehen von Institutionen stärken und gesellschaftliche Zustimmung sichern. Dabei zeigen sich in der bildungspolitischen KI-Debatte zwei zentrale Rechtfertigungsmuster: (1) KI als Instrument des Reputationsgewinns und (2) das Streben nach Akzeptanz als Realitätsprüfung der Cité der Meinung.

9.3.5.1. Künstliche Intelligenz als Instrument des Reputationsgewinns

Die zentrale Rechtfertigung der Cité der Meinung liegt in der strategischen Nutzung von KI zur Imagebildung. Der Einsatz innovativer Technologien soll die öffentliche Wahrnehmung von Hochschulen, Initiativen und staatlichen Programmen verbessern. So wird der KI-Einsatz etwa als Instrument zur Anwerbung von Studierenden genutzt: „Das Spiel mit künstlerischer

Intelligenz (…) begeistert Jahr für Jahr immer mehr junge Menschen und bildet damit für die Universität eine Plattform zur Werbung von Studierenden in den MINT-Fächern“ (GWK_1045, 2019, S. 109). KI wird in diesem Fall als aufmerksamkeitswirksames Symbol technologischer Zukunftsfähigkeit betrachtet. Dabei handelt es sich um eine Reputationslogik, die auch auf nationaler Ebene zum Ausdruck kommt: Die Bundesregierung betont das internationale Renommee deutscher KI-Forschungslandschaften, insbesondere des Cyber Valley, welches es weiter zu stärken gelte (BT_162b, 2019, S. 11; GWK_1051, 2020, S. 396). Gleichzeitig wird auch unter Bezug auf die Cité der Meinung auf die Reputation internationaler Vorbilder wie Finnlands KI-Bildungsprogramm oder Estlands digitale Verwaltung als Benchmark verwiesen, an denen sich deutsche Initiativen messen lassen müssten (BT_113b, 2020, S. 199; LT-SH_934, 2021, S. 25; TAB_1325, 2022, S. 74). Ebenso werden Hochschulen teils positiv hinsichtlich ihres KI-Einsatzes hervorgehoben (TAB_1335, 2020, S. 66), wenngleich auf bestehende Diskrepanzen zwischen Anspruch und Realität hingewiesen wird: So bliebe die Umsetzung an Hochschulen bislang hinter den selbst zugeschriebenen Bedeutungen zurück (EFI_1082, 2019, S. 16). Die Cité der Meinung operiert damit im Modus reputativer Orientierung an internen und externen Vorbildern und in der Hoffnung, selbst Anerkennung zu generieren.

9.3.5.2. Akzeptanz als notwendige Voraussetzung des Einsatzes Künstlicher Intelligenz

Eng mit der Reputation verknüpft ist das Streben nach breiter öffentlicher Akzeptanz, welches durch den überwiegend technikoptimistischen Tenor der bildungspolitischen KI-Debatte geprägt wird. Die FDP fordert beispielsweise „eine breite Akzeptanz künstlicher Intelligenz als Lernhilfsmittel“ (BT_288a, 2020, S. 4) sowie eine gezielte „Aufklärungskampagne über die Chancen der KI“ (BT_157b, 2019, S. 7). Ebenso betonen verschiedene Landesparlamente, dass KI nur dann wirksam eingesetzt werden könne, wenn die Technologie gesellschaftlich akzeptiert werde (LT-SH_592, 2019, S. 5; LT-N_798, 2021, S. 45; Werner/Die Linke, LT-TH_850, 2021, S. 3540; LT-SH_897, 2021, S. 160), denn davon sei etwa die Bereitschaft von Lernenden zur Freigabe ihrer persönlichen Daten abhängig (TAB_1333, 2021, S. 4). Trotz der positiven Technikdeutung wird das Streben nach Akzeptanz zu einer zentralen Realitätsprüfung der Cité der Meinung. Das TAB hebt „sowohl ein Mangel politischer Unterstützung als auch eine negative Grundhaltung (…) bei Entscheidungsträgern, Eltern und Lehrern“ (TAB_1335, 2020, S. 83) kritisch hervor. So wird ein Mangel an Zustimmung nicht nur als Implementationshindernis (HFD_1187, 2015, S. 51; HFD_1202, 2021, S. 17), sondern ebenso als Reputationsschaden erachtet, der sich in Widerstand und damit schließlich in einem Scheitern der KI-Anwendung niederschlagen kann (HFD_1202, 2021, S. 48). Dieses Risiko wird anhand prominenter internationaler Fehlschläge illustriert, wie das von der Bill & Melinda Gates Foundation mit

100 Millionen Dollar geförderte Learning-Analytics-Projekt, das aufgrund massiver Kritik von Eltern- und Lehrendenverbänden eingestellt wurde (BT_244b, 2020, S. 7; Völlers/SPD, BT_244c, 2019, S. 14626). In ähnlicher Weise wird der britische KI-Abiturnotenskandal als Beleg dafür angeführt, dass „eine übereilte und schlecht vorbereitete Einführung von [Learning Analytics-]Systemen deren Einsatz scheitern lassen kann" (TAB_1333, 2021, S. 5). Diese Beispiele fungieren als warnende Realitätsprüfungen, die verdeutlichen, dass Akzeptanzverlust politisch wirkmächtig wird, technologische Projekte delegitimieren kann und politischem Ansehen schaden kann.

Darüber hinaus weiten sich die Realitätsprüfungen der Cité der Meinung auf die Gestaltung öffentlicher Diskurse aus. Die Enquete-Kommission mahnt, dass „Chancen und Risiken des KI-Einsatzes (…) emotional und inhaltlich überhöht [werden]. Beides wird dadurch unsachgemäß verstärkt" (BT_113b, 2020, S. 348). Dies zeige sich etwa anhand der öffentlichen Debatte über ChatGPT, die zwischen „utopischer Naivität" und „dystopischer Panik" schwanke (HFD_1317, 2022, S. 2; s. a. HFD_1202, 2021, S. 48). Vor dem Hintergrund des hohen Technikoptimismus kritisiert Grünen-Abgeordnete Stumpp den „KI-Hype", dem insbesondere die FDP verfallen sei (BT_244c, 2019, S. 14551). Ein übersteigertes Erwartungsniveau berge das Risiko von Vertrauens- und Imageverlusten, wenn Systeme den öffentlich zugeschriebenen Erwartungen nicht gerecht werden könnten (HFD_1202, 2021, S. 24). Am Beispiel der KI-basierten Emotionserkennung hebt das TAB die Gefahr dieses Vertrauensverlusts hervor: „Ein Schaden für den öffentlichen Sektor könnte sich dann ergeben, wenn Fördermittel in Anwendungsforschung zur Entwicklung marktfähiger KI-Systeme zur Emotionserkennung fließen, bevor die wissenschaftlichen Grundlagen für die Entwicklung valide arbeitender Systeme gelegt wurden" (TAB_1327, 2021, S. 6). Damit rückt die Cité der Meinung das Risiko überhöhter Erwartungen und vorschneller Implementierungen vor allem als Gefahr für das institutionelle Ansehen in den Vordergrund, während mögliche Nachteile für die Betroffenen der Systeme in den Hintergrund treten.

Um Akzeptanz langfristig zu stärken und der Realitätsprüfung somit Rechnung zu tragen, verfolgen die Akteur:innen der bildungspolitischen Debatte drei komplementäre Strategien: (1) Wissensförderung, (2) Stärkung des öffentlichen Diskurses und (3) menschenzentrierte KI-Entwicklung. Wissenschaftliche Erkenntnisse etwa zu erklärbarer KI oder zur Mensch-Maschine-Interaktion sollen die KI-Entwicklung an Akzeptanzfaktoren ausrichten (TAB_1324, 2022, S. 4; HFD_1202, 2021, S. 48) und zugleich „zu einer Versachlichung der Debatte" (HFD_1202, 2021, S. 48) beitragen. Gesamtgesellschaftliche Bildungsangebote sollen zudem helfen, Berührungsängste abzubauen und das Verständnis und die Akzeptanz von KI zu fördern (BT_113b,

2020, S. 199; GWK_1044, 2018, S. 328; LT-SH_670, 2019, S. 14; LT-SH_934, 2021, S. 25). Darüber hinaus gelte es, den öffentlichen Diskurs zu stärken: „Nur über einen intensiven gesellschaftlichen Dialog kann es breite Akzeptanz für und Neugierde auf KI geben" (Harms/SSW, LT-SH_970, 2022, S. 10726; vgl. auch BT_113b, 2020, S. 317; Röspel/SPD, BT_113d, 2020, S. 23762; FDP, BT_244b, 2020, S. 5; CDU, LT-BW_793, 2020, S. 6; LT_SH_934, 2021, S. 11). Meyer (CDU) fordert explizit eine Politisierung der KI-Thematik: „Wir wollten das (…) noch stärker in die Öffentlichkeit tragen, denn es muss Akzeptanz in der Bevölkerung finden. Wir müssen die Leute mitnehmen. Dazu gehört, dass wir über solche Themen im politischen Raum sprechen" (LT-SN_904, 2021, S. 2793). Schließlich gilt eine wertebasierte KI-Entwicklung als Voraussetzung für Vertrauen und langfristige gesellschaftliche Zustimmung (LT-SH_897, 2021, S. 153), wobei etwa „Trusted Learning Analytics" das Renommee einsetzender Hochschulen erhöhen und zugleich Akzeptanz fördern sollen (LT-HE_765, 2020, S. 7).

Aus Sicht der Cité der Meinung ist der KI-Einsatz im Bildungsbereich somit vor allem eine Frage reputativer Wirkung und öffentlicher Zustimmung, die über Erfolg oder Scheitern entscheidet. Dabei zeigt sich ein Spannungsverhältnis zwischen beabsichtigter Politisierung zur Akzeptanzförderung und der Gefahr unerwünschter Politisierung bei gesellschaftlichem Widerstand oder überhöhten Erwartungen.

9.4. Zur Rolle solutionistischer Rechtfertigungen des Einsatzes Künstlicher Intelligenz

Im nächsten Schritt der Ergebnisdarstellung rücken die Rechtfertigungen aus der Perspektive der Cité der Solution in den Mittelpunkt. Vor dem Hintergrund des in der Analyse bereits implizit zum Ausdruck gekommenen techniksoziologischen Spannungsverhältnisses zwischen Gestaltbarkeit und technologischer Zwangsläufigkeit sowie der wachsenden wissenschaftlichen Kritik an solutionistischen Technikdeutungen (siehe Abschnitt 6.4.4) wird in diesem Abschnitt die zweite Forschungsfrage adressiert und in den Blick genommen, inwiefern sich solutionistische Rechtfertigungen in der bildungspolitischen Debatte über die KI-Einführung manifestieren. Dabei zeigt sich, dass solutionistische Rechtfertigungen in mehr als jedem sechsten codierten Segment auftreten und nahezu ausschließlich positiv konnotiert sind. Sie stützen sich überwiegend auf Wahrheitsprüfungen, in denen KI als notwendiger Bestandteil zukünftiger Bildungs- und Gesellschaftsentwicklung dargestellt wird. Die folgende Analyse fokussiert auf zentrale Muster dieser Rechtfertigungen: erstens die Annahme der Unvermeidbarkeit von KI im Sinne eines optimistischen Determinismus, zweitens die Darstellung von KI als Problemlösungsinstrument und drittens die Diffusität solutionistischer Nutzenversprechen. Anschließend

werden Realitätsprüfungen rekonstruiert, die die tatsächliche Problemlösungsfähigkeit hinterfragen. Darüber hinaus werden fehlende Rechtfertigungen als Ausdruck eines impliziten Solutionismus rekonstruiert, wodurch der Anteil solutionistisch geprägter Rechtfertigungen an der Gesamtzahl der Segmente auf 30 % steigt. Abschließend werden Kompromisse mit der Cité der Industrie und der Cité des Staatsbürgers betrachtet, die die instrumentelle Einbindung der Cités in solutionistische Rechtfertigungen verdeutlichen.

9.4.1. Die Unvermeidbarkeit Künstlicher Intelligenz als Ausdruck des optimistischen Determinismus solutionistischer Rechtfertigungen

Als zentrales Element der solutionistischen Rechtfertigung rekonstruiert die Analyse eine in der bildungspolitischen Debatte verbreitete deterministische Annahme, dass KI den gesellschaftlichen Wandel unausweichlich prägt und somit „eine unumstößliche Tatsache“ (Hartung/SPD, LT-TH_850, 2021, S. 3545) darstelle. In besonders zugespitzter Form formuliert der AfD-Abgeordnete Jongen eine schicksalhafte Zwangsläufigkeit in Referenz des Philosophen Seneca: „Man könnte auch sagen: KI ist das Schicksal. Und wie schon Seneca wusste: Den Willigen führt das Schicksal, den Unwilligen treibt es“ (BT_113d, 2020, S. 23762). Gerade in Parlamentsdebatten und politischen Strategiepapieren sprechen politische Entscheidungsträger:innen KI entsprechend das Potenzial zu, „die Digitalisierung insgesamt auf eine neue Ebene zu heben“ (LT-SH_670, 2019, S. 55). KI gilt als tiefgreifende Kraft, die gesellschaftliche Prozesse in Bereichen des Arbeitens, der Bildung, der Wirtschaft oder des Alltags zum Positiven verändere (LT-SH_592, 2019, S. 2, Althusmann/CDU, LT-NI_776, 2020, S. 8600). Die Zukunft erscheint in diesen Darstellungen bereits vorgezeichnet und die politische Aufgabe besteht darin, den Wandel voranzutreiben. Somit begründet diese wahrgenommene Unausweichlichkeit eine Einengung der politischen Debatte weg von der grundsätzlichen Frage der KI-Einführung hin zur Frage der Art und Weise des KI-Einsatzes, wie ein Blogbeitrag des HFD erfasst:

> Das Rad lässt sich nicht zurückdrehen. KI-Sprachmodelle und -systeme sind ein Faktum im Bereich der Wissensarbeit. (Hoch-)Schulseitige Vogel-Strauß-Taktik ist daher fehl am Platz. Wenn es aber nicht mehr um das ‚Ob‘ der Verwendung von KI-Tools geht, dann muss die Frage lauten: Wie soll der Einsatz künftig gestaltet werden? (HFD_1317, 2022, S. 5)

Die politische Antwort auf diese Frage lautet einer solutionistisch-deterministischen Rechtfertigungslogik folgend mitunter flächendeckende Investitionen und Maßnahmenpakte der KI-Entwicklung auf den Weg zu bringen, die als Investition in die Zukunft erachtet werden und „deren Ergebnisse erst in den nächsten Jahren und Jahrzehnten sichtbar und Wertschöpfung im

Land generieren werden“ (LT-SH_592, 2019, S. 2), wie einem Gesetzentwurf des Landes Schleswig-Holstein zu entnehmen ist und wie Plambeck (CDU) ebenso hervorhebt:

> Dabei wird in absehbarer Zeit fast jeder Lebensbereich durch KI beeinflusst, sei es die Medizin, die Bildung, die Mobilität oder die Wirtschaft. Das Schöne ist, dass diese Zukunft bereits heute in Schleswig-Holstein gelebt wird – insbesondere dank einer innovativen Jamaika-Koalition, die 2019 zu Recht auf dieses Thema gesetzt hat. (LT-SH_970, 2022, S. 10721)

Diese Äußerung ist somit zugleich ein gutes Beispiel dafür, wie gerade Regierungsparteien Wahrheitsprüfungen zur Rechtfertigung ihrer Politik einsetzen, indem hier die deterministisch-solutionistische Grundannahme der Unausweichlichkeit von KI als Rechtfertigung politischen Handelns herangezogen wird. Ähnlich argumentiert Brandenburg (FDP), der eine Förderung KI-basierter Lernmethoden als eine an die „Lebenswirklichkeit der Menschen“ (BT_251f, 2020, S. 18394) angepasste Notwendigkeit darstellt und deren Einsatz als natürliche Konsequenz des digitalen Alltags rahmt. Diese solutionistische Rechtfertigungslogik schränkt somit nicht nur das Kontingenzbewusstsein politischen Handelns auf die Frage der Ausgestaltung des KI-Einsatzes ein, sondern lässt die Investition in KI gleichzeitig als Notwendigkeit erscheinen.

9.4.2. Künstliche Intelligenz als Problemlösungsinstrument und Verbesserung des Status quo

Den Kern solutionistischer Rechtfertigungen bildet die Annahme, KI sei ein universell einsetzbares Werkzeug politischer Problemlösung. So betont Christmann (Die Grünen): „Künstliche Intelligenz ist keine Supermacht, die einfach über uns kommt, sondern künstliche Intelligenz ist eine Technologie, die uns ganz konkret bei konkreten Herausforderungen helfen kann“ (BT_157c, 2018, S. 8096). Damit wird zwar eine politische Steuerbarkeit suggeriert (und somit ein strenger Determinismus in die Schranken gewiesen), die Problemlösungskapazität der KI jedoch trotz des Verweises auf konkrete Herausforderungen zugleich breit adressiert. Um das mannigfaltige Potenzial der KI zu unterstreichen, wird im bildungspolitischen Diskurs dabei auf erfolgreiche KI-Anwendungen in anderen gesellschaftlichen Bereichen verwiesen, wie in der Medizin oder im Bereich der Bekämpfung des Klimawandels und der Corona-Pandemie (Christmann/Die Grünen, BT_113d, 2020, S. 23761; Waldinger-Thiering/SSW, LT-SH_830, 2021, S. 8748; Hartung/SPD, LT-TH_850, 2021, S. 3545; Meyer/CDU, LT-SN_904, 2021, S. 2793; LT-SH_934, 2021, S. 11; WR_1175, 2022, S. 21). Der Rückgriff auf diese Beispiele fungiert als Wahrheitsprüfung, indem die Funktionsfähigkeit der Technologie in anderen Feldern als Evidenz für ihren Nutzen im Bildungssystem herangezogen wird. Das Potenzial der KI wird somit nahezu grenzenlos eingeschätzt, weshalb Investitionsstrategien die Förderung von

Forschung und Anwendungen in nahezu allen gesellschaftlichen Bereichen vorsehen (BT_162b, 2019, S. 6; LT-NW_895, 2021, S. 143; LT-SH_934, 2021, S. 3).

Während die Universalität der Einsatzmöglichkeiten häufig in diffusen Vorstellungen von Verbesserungspotenzialen mündet (siehe vertiefend dazu Abschnitt 9.4.3), liefert der Antrag der FDP-Fraktion „Alphabetisierung durch Künstliche Intelligenz – Chance für jeden" eines der eher seltenen Beispiele für die Verwendung von KI zur Lösung konkreter identifizierter Bildungsprobleme. Kritisiert wird, „dass die Bundesregierung nicht konsequent den Einsatz Künstlicher Intelligenz (KI) im Bildungsbereich fördert und somit einen Großteil an Potenzial ungenutzt lässt" (BT_288a, 2020, S. 2). Dabei könnte durch den Einsatz von KI „in der Grundbildung einer geringen Literalisierung entgegengewirkt werden" (BT_288b, 2021, S. 2). Der Antrag illustriert damit den Kern solutionistischer Rechtfertigungen, indem KI als wirksames und zugleich einzig sinnvolles Instrument zur Lösung eines dargelegten Problems dargestellt wird.

Häufiger als konkrete Problemlösungsvorschläge finden sich solutionistische Rechtfertigungen jedoch in der Auflistung potenzieller, meist vage benannter Einsatzbereiche. So formuliert Plambeck (CDU) weitgehend unspezifisch, KI könne „Kompetenzen wie Lesen, Schreiben, Rechnen, aber auch das begleitete Lernen in der Erwachsenenbildung (…) erheblich unterstützen" (LT-SH_830, 2021, S. 8743). Auch automatisierte Bewertungssysteme, Chatbots oder KI-basierte Motivationsanalysen werden in solutionistischen Rechtfertigungen als wünschenswert dargelegt, jedoch ohne konkrete Bedarfe oder alternative Perspektiven darzulegen (HFD_1201, 2021, S. 2; WR_1130, 2019, S. 75). Gleiches gilt für den administrativen Anwendungsbereich, in dem die nicht näher spezifizierten „vielfältigen Einsatzmöglichkeiten und das sich daraus ergebende Potential" (LT-SH_897, 2021, S. 69) genutzt werden sollten, da KI „den nächsten Schritt in der Weiterentwicklung moderner Informationstechnologie" (ebd.) darstelle. Während der KI-Einsatz an dieser Stelle somit als scheinbar unhinterfragte natürliche Entwicklung erscheint, mündet diese vereinzelt in Rechtfertigungen eines „second-order solutionism" (Nachtwey & Seidl, 2023), wenn etwa Plagiatserkennungssysteme als Reaktion auf den verstärkten Einsatz generativer KI gefordert werden (HFD_1200, 2020, S. 5). Damit wird eine technische Lösung als Antwort auf eine KI-induzierte Herausforderung präsentiert, was die selbstverstärkende Logik solutionistischer Rechtfertigungen verdeutlicht.

KI wird in diesem Kontext als *das* zentrale Instrument zur grundlegenden Verbesserung des Status quo[33] des Bildungssystems dargestellt, wie Suding (FDP) deutlich macht:

> Wenn wir Learning Analytics an unseren Schulen fördern und dem einen klaren Rahmen setzen, dann stärken wir die Bildung unserer Kinder und dann schützen wir gleichzeitig ihre Lerndaten. Das und nicht weniger haben unsere Schülerinnen und Schüler verdient: weltbeste Bildung. (BT_244c, 2019, S. 14547)

Wenngleich in diesem Beitrag die Herausforderung des Datenschutzes anerkannt wird (siehe dazu Abschnitt 9.5.1.3), ist aus Sicht der FDP klar, dass KI positive Effekte erzeugt. Auch die Bund-Länder-Vereinbarung „Künstliche Intelligenz in der Hochschulbildung" führt eine Vielzahl möglicher Anwendungen an, die pauschal als „Verbesserung der Hochschulbildung durch den Einsatz von KI" (GWK_1052, 2020, S. 2) gerahmt werden, wie zum Beispiel KI-gestützte Studienberatung, Lernprozessunterstützung oder Modulentwicklung. Dabei verfolgt die Vereinbarung das Ziel, KI „als Schlüsseltechnologie in der Breite des Hochschulsystems zu verankern" (WR_1177, 2022, S. 49). Auf Basis dieser solutionistischen Rechtfertigungen treiben politische Akteur:innen die KI-Implementierung voran, ohne konkrete Belege für das Eintreten der zugesprochenen Potenziale vorzulegen oder unintendierte Folgen zu berücksichtigen.

9.4.3. Die Diffusität solutionistischer Rechtfertigungen als Ausdruck eines digitalen Geists

Wie die unspezifische Auflistung verheißungsvoller Anwendungsbereiche bereits nahelegt, sind solutionistische Rechtfertigungen häufig durch eine diffuse Vorstellung von Verbesserung geprägt. Der wahrgenommene Veränderungsbedarf wird nicht, wie etwa in der Cité der Industrie, an klaren Zielgrößen wie Effizienzgewinnen festgemacht, sondern bleibt in vielen Fällen unspezifisch. Investitionen werden beispielsweise „[z]ur systematischen Verbesserung der Lehr- und Studienbedingungen" (GWK_1038, 2015, S. 54) gerechtfertigt, ohne Indikatoren für eine nachweisbare Verbesserung zu benennen. Besonders deutlich wird diese Unspezifiziertheit in der Aussage des hessischen Kultusministers Lorz (CDU): „Meine Damen und Herren, selbstverständlich wollen wir auch die Potenziale der KI – was immer das jetzt ist, das mögen Sie beide unter sich ausmachen – für hessische Schulen nutzbar machen" (LT-HE_1011, 2022, S. 8594). Es bleibt somit unklar, welche Probleme durch den KI-Einsatz gelöst oder welche Rahmenbedingungen verbessert werden sollen, solange die Lösung eine technologische ist.

[33] Die auf eine Verbesserung des Status quo zielenden Rechtfertigungen der Cité der Solution weisen Parallelen zum Optimierungsstreben der Cité der Industrie auf (siehe Abschnitt 9.3.1.1). Der zentrale Unterschied liegt jedoch in der Zielrichtung der Argumentation: Während solutionistische Versprechungen häufig unspezifisch bleiben und vage auf eine allgemeine Verbesserung gesellschaftlicher Zustände verweisen, zielen industrielle Rechtfertigungen explizit auf die Steigerung von Leistung, Effizienz und Produktivität ab.

So manifestiert sich an dieser Stelle die deterministische Grundannahme, die KI als inhärent wirksames Mittel konzipiert, das nahezu beliebig auf unterschiedliche Problemfelder übertragbar erscheint. Die noch von Christmann (Die Grünen) benannten „konkreten Herausforderungen“ (BT_157c, 2018, S. 8096) bleiben daher häufig ohne Substanz. Die Rechtfertigungslogik folgt vielmehr dem von Nachtwey und Seidl (2023) beschriebenen digitalen Geist, der die Legitimationsvorstellung stützt, dass der Einsatz von KI per se sinnvoll sei, unabhängig von der spezifischen Problemlage. Ein weiteres Beispiel hierfür liefert die Strategie des Landes Schleswig-Holstein, die betont: KI ist „kein Selbstzweck, sondern sie trägt dazu bei, politische, wirtschaftliche und gesellschaftliche Ziele besser zu erreichen. Hierzu gehören die Sicherung des Wohlstandes, die Wettbewerbsfähigkeit der schleswig-holsteinischen Wirtschaft, bürgerfreundliche Politik und Verwaltung sowie die nachhaltige Entwicklung des Landes“ (LT-SH_934, 2021, S. 10). Damit wird sichtbar, dass KI nicht nur als singuläre Lösung, sondern als Technologie verstanden wird, die mehrere Zielsetzungen gleichzeitig erfüllen soll und somit eine Art Querschnittslogik symbolisiert. Insofern können Bezüge zu anderen Cités in solutionistischen Rechtfertigungen zwar aufgegriffen werden, sie bleiben jedoch unscharf und in der übergeordneten Logik des Solutionismus eingebettet, der im metaphorischen Sinn beabsichtigt, mehrere Fliegen mit einer Klappe zu schlagen. Diese Diffusität, die sich in einem optimistischen Determinismus und der bewusst offengehaltenen Breite potenzieller Einsatzfelder zeigt, ist charakteristisch für solutionistische Rechtfertigungen. KI erscheint dabei als Technologie, die unabhängig von Kontext oder Problemlage zur Verbesserung des Status quo beitragen könne.

9.4.4. Der innere Widerstreit des Solutionismus: eine Realitätsprüfung der Problemlösungsfähigkeit und Wirkkraft Künstlicher Intelligenz

Während die zuvor rekonstruierten solutionistischen Rechtfertigungen vor allem Ausdruck selbstbestätigender Wahrheitsprüfungen sind, lassen sich innerhalb der Cité der Solution auch Spannungen erkennen, die in Form von Realitätsprüfungen die tatsächliche Problemlösungsfähigkeit von KI hinterfragen. Kolbe (SPD) hebt dies im Vorwort des Enquete-Berichts hervor: KI „kann ein Werkzeug sein, um Probleme zu lösen oder sie womöglich auch zu verschlimmern“ (BT_113b, 2020, S. 27). Der Enquete-Bericht betont daher die Grenzen eines unkritischen Technologieglaubens und hält fest,

> dass KI-Systeme ein mächtiges Werkzeug, aber eben nur ein Werkzeug sind. Parlament und Regierung müssen weiterhin politische Lösungen für gesellschaftliche Herausforderungen finden – KI kann dann für die Umsetzung eingesetzt werden. Mitunter öffnen sich

> für gesellschaftliche Herausforderungen aber auch neue Lösungsansätze durch KI. (BT_113b, 2020, S. 31 f.)

Trotz der Betonung politischer Handlungsfähigkeit wird KI somit an dieser Stelle ein Problemlösungspotenzial attestiert. Demgegenüber befürchten einige politische Akteur:innen jedoch, dass bestehende gesellschaftliche Konflikte nicht durch KI gelöst, sondern vielmehr verschärft würden: „die Konflikte, die wir schon vor der Digitalisierung und vor KI hatten, [werden sich] nicht einfach durch KI lösen, sondern sie werden sich erneut vertiefen“ (Sitte/Die Linke, BT_113d, 2020, S. 23760). Die AfD kritisiert entsprechend eine bildungspolitische Digitalisierungsagenda, die „dem Motto ‚neu ist immer besser‘“ (Jankowski/AfD, LT-TH_850, 2021, S. 3544) folge und ohne pädagogisch fundierte Bedarfsprüfung verlaufe. Gerade bei Grundschüler:innen zeige die Forschung, „dass die Nachteile der Digitalisierung deutlich größer sind als die vermeintlichen Vorteile“ (ebd.). Demnach beruhe das Streben nach einer umfassenden Digitalisierung und dem Einsatz von KI im Bildungswesen auf einem grundlegenden Trugschluss, da es einem „verkürzte[n] Bildungsbegriff“ (Kaufmann/AfD, LT-TH_736, 2020, S. 1522) folge, der die nachhaltige Verinnerlichung von Lerninhalten vernachlässige.

Diese Kritik richtet sich insbesondere gegen technikoptimistische Positionen der FDP, die CDU und Grünen hinterfragt werden: „Die FDP bleibe sich treu und beschwöre Digitalisierung und Künstliche Intelligenz wieder einmal als Allheilmittel im Bildungsbereich“ (CDU, BT_288b, 2021, S. 11). Auch die Abgeordnete Stumpp (Die Grünen) schließt sich dieser Kritik an, indem sie die Anträge der FDP-Fraktion als inhaltlich beliebig und als Ausdruck einer Symbolpolitik beschreibt: „Beim Lesen Ihres Antrags drängte sich mir zunehmend der Eindruck auf, Digitalisierung – gestern in aller Munde – reicht heute nicht mehr. Das neue Buzzword ist KI, und deswegen muss KI diese Woche in jedem Antrag der FDP vorkommen, Sinnhaftigkeit second“ (BT_244c, 2019, S. 14551). Allerdings relativiert die FDP vereinzelt auch selbst überhöhte solutionistische Erwartungen. Brandenburg (FDP) betont, es sei falsch, „gesellschaftliche Probleme oder politische Wünsche auf eine Technologie outzusourcen“ (BT_113d, 2020, S. 23759), da KI lediglich ein Werkzeug bleibe, das „die Probleme, die wir als Menschheit über Tausende von Jahren aufgebaut haben“ (ebd.), nicht allein lösen könne. Während der Abgeordnete dennoch das Potenzial von KI hervorhebt, zukünftige Entwicklungen positiv zu beeinflussen, kann diese Einordnung als bewusstes Erwartungsmanagement interpretiert werden, um übersteigerte Erwartungen frühzeitig zu dämpfen und späteren Enttäuschungen vorzubeugen, die den Einsatz von KI diskreditieren könnten.

In ähnlicher Weise unterstreicht das TAB, dass sich digitale Medien „nicht per se (positiv oder negativ) auf Lehr- und Lernmethoden“ (TAB_1335, 2020, S. 242) auswirkten. Damit rückt zunehmend die Frage in den Vordergrund, welche Probleme durch KI tatsächlich adressiert werden können. Der Enquete-Bericht verweist darauf, dass die Stärken der KI in Bereichen analytischer Verarbeitung lägen, die Leistungsfähigkeit von KI jedoch bei Kreativität und kritischem Denken hinter dem Menschen zurückbleibe (BT_113b, 2020, S. 311 f., 319). Diese Relativierungen machen den inneren Spannungsbogen der solutionistischen Realitätsprüfung sichtbar: Einerseits wird KI als technologisches Problemlösungsinstrument dargestellt, andererseits wird ihre zugesprochene Wirkungskraft immer wieder begrenzt.

Dieses Spannungsfeld manifestiert sich darüber hinaus auch in der Bewertung des deterministischen Einflusses von KI. So betont etwa Brünler (Die Linke):

> KI dringt unaufhaltsam in privates und berufliches Leben vor und das nicht immer nur zum Vorteil. Denken wir nur an das Diskriminierungspotenzial bestimmter, oftmals nur scheinbar objektiver Algorithmen; denn diese reproduzieren in der Regel oft die subjektiven Wertmaßstäbe oder Vorurteile ihrer Architekten. Dennoch: KI verändert zunehmend die Art und Weise, wie wir kommunizieren, produzieren und konsumieren, aber auch, wie wir lernen und studieren. (LT-SN_904, 2021, S. 2797 f.)

Der Abgeordnete verweist damit auf Ambivalenzen, wie sie seit Jahrzehnten in der STS-Literatur diskutiert werden: Die politische Debatte über KI bewegt sich einerseits zwischen der Vorstellung einer technikgetriebenen, unausweichlichen Entwicklung und der Betonung ihrer sozialen Gestaltbarkeit, andererseits zwischen technikoptimistischen und -pessimistischen Positionen. Dabei identifiziert Schön (Union) eine Polarisierung der öffentlichen Debatte, in der die zunehmende Leistungsfähigkeit von KI gleichermaßen utopische wie dystopische Projektionen befeuere:

> Das beflügelt Fantasien. Die einen denken an den Terminator, sehen die Maschine, die die Macht über die Menschen übernimmt. Die anderen setzen eine Menge Hoffnungen in KI (…). Sie sehen darin die Lösung für viele Probleme der Menschheit (…). Alles scheint möglich mit künstlicher Intelligenz. Deshalb ist es kein Hype, über den plötzlich alle sprechen, weil es gerade modern ist; es ist Fakt, dass KI derzeit der größte Treiber der Digitalisierung ist und nahezu alle Bereiche (…) beeinflusst und mitgestaltet. (BT_113c, 2018, S. 4152)

Die Realitätsprüfung der Cité der Solution richtet damit den Blick auf die zentrale Frage, ob KI die ihr zugeschriebenen Problemlösungskompetenzen tatsächlich einlösen kann. Zwischen

hohen Erwartungen an die transformative Kraft der Technologie und gleichzeitigen Mahnungen zu realistischen Einschätzungen entsteht ein Spannungsfeld, das solutionistische Rechtfertigungen in der bildungspolitischen Debatte herausfordert und ihre Grenzen offenlegt.

9.4.5. Mangelnde Rechtfertigung als Ausdruck eines deterministischen Optimismus: Der implizite Solutionismus

Neben den verschiedenen Cités, die als moralische Grundlage für den Einsatz von KI herangezogen werden, offenbart die Analyse auch eine Reihe von Forderungen, denen eine explizite Rechtfertigung fehlt. Dies betrifft knapp 13 % der codierten Segmente. In diesen Beiträgen wird KI in der Regel ein hoher Stellenwert zugesprochen, dieser wird jedoch vielfach nicht weiter gerechtfertigt. So betont etwa Thüringens Minister für Wirtschaft, Wissenschaft und Digitale Gesellschaft Tiefensee (SPD), „dass wir uns mehr um die Künstliche Intelligenz kümmern sollten, [sei] sehr wichtig“ (LT-TH_736, 2020, S. 1500), ohne die zugrunde liegende Bedeutung näher zu begründen. Mangelnde Rechtfertigungen finden sich dabei insbesondere im Kontext (1) der Förderung von KI und (2) des Wissenserwerbs über KI.

9.4.5.1. Mangelnde Rechtfertigung der Investitionen in den Aufbau Künstlicher Intelligenz

Häufig beziehen sich nicht weiter gerechtfertigte Äußerungen auf Berichte über bereits getätigte Investitionen in KI, implementierte KI-Anwendungen oder initiierte Maßnahmen (GWK_1047, 2019, S. 395; LT-NI_568, 2019, S. 2; LT-SH_592, 2019, S. 8; LT-HE_765, 2020, S. 7; LT-NI_798, 2021, S. 24 f.; Günther/CDU, LT-SH_830, 2021, S. 8743; LT-NW_895, 2021, S. 88; Fritzsche/CDU, LT-SN_904, 2021, S. 2790; Lorz/CDU, LT-HE_1011, 2022, S. 8593; WR_1158, 2021, S. 10 f.) und bilden so Paradebeispiele der Wahrheitsprüfungen. Regierungen berichten etwa über den Aufbau von Kompetenzzentren (BT_162b, 2019, S. 10), Fortschritte im DigitalPakt Schule (Kaczmarek/SPD, BT_244d, 2021, S. 27064) oder darüber, dass der „Einsatz von Künstlicher Intelligenz in der digitalen Lehre sowie eine landesweite Infrastruktur für Virtualisierungskonzepte digitaler Labore“ (LT-NW_895, 2021, S. 89) gefördert werde, ohne Rechtfertigungen für diese Maßnahmen darzulegen. Dabei lässt sich die mangelnde Rechtfertigung teilweise durch spezifische Dokumententypen erklären, etwa durch Regierungsantworten auf Faktenabfragen im Rahmen parlamentarischer Anfragen. Gleichzeitig kann nicht ausgeschlossen werden, dass getätigte oder geplante Investitionen in KI mitunter bereits Gegenstand vorangegangener Debatten waren, was es an dieser Stelle limitierend zu berücksichtigen gilt (siehe dazu Abschnitt 11.2). Gleichwohl fehlt auch in Berichten über zukünftige Investitionen in KI und die digitale Bildungsinfrastruktur vielfach eine explizite Rechtfertigung (BT_157b, 2019, S. 5 f.; GWK_1038, 2015, S. 33 f.; LT-TH_836, 2021, S. 8 f.; TAB_1334, 2021, S. 16). So heißt es beispielsweise, es sei „zu prüfen, inwieweit künstliche

Intelligenz zukünftig zur Unterstützung der Lehrkräfte in den Unterricht einbezogen werden könnte“ (LT-NI_662, 2019, S. 4), ohne den Einsatz von KI grundsätzlich zu rechtfertigen. Auch die Forderung der FDP-Fraktion nach einem „Digitalpakt 2.0“ mit umfassenden Investitionen in die digitale Infrastruktur sowie einer „Forschungsförderung in den Bereichen Learning Analytics und Educational Data Mining“ (BT_195b, 2019, S. 5) bleibt ohne fundierte Rechtfertigung. Ähnlich wird die Nutzung bestehender Datenbestände durch KI-Anwendungen begrüßt, ohne dass ein klar umrissenes Ziel benannt würde (WR_1175, 2022, S. 30).

Zwar wird die Nutzung von KI und die damit verbundene Förderung nicht immer völlig voraussetzungslos vorangetrieben, da mitunter Evaluationsmaßnahmen beabsichtigt sind. Auffällig ist jedoch, dass diese Evaluationskonzepte fast ausschließlich auf eine Ausweitung der Investitionen ausgerichtet sind. So gelte es dem WR zufolge zu prüfen, „ob mit den Fördervorhaben etwa im Bereich Big Data und Künstliche Intelligenz die Förderung ausreichend ausgestaltet wurde“ (WR_1143, 2020, S. 58) oder wo zusätzliche Bedarfe bestünden. Restriktive Alternativen oder kritische Optionen werden hier nicht adressiert. Ebenso bleiben Empfehlungen einer verstärkten Auseinandersetzung der Hochschulen mit innovativen Lehr-Lern-Modellen (EFI_1078, 2015, S. 57) oder Forderungen nach einem „umfassenden Ausbau der Nutzung und Entwicklung von KI im Bildungsbereich sowie in der Erwachsenengrundbildung“ (BT_288b, 2021, S. 6) ohne ergänzende Rechtfertigung. Auch der Appell Sudings (FDP), Lehrkräfte zur Nutzung von KI-Systemen zu ermutigen, verdeutlicht diesen investitionsorientierten Argumentationsmodus, der auf eine inhaltlich tragfähige Rechtfertigung verzichtet:

> Wir müssen darüber hinaus die Lehrkräfte unterstützen; denn mit ihnen steht und fällt die Nutzung von Lerndaten. Einige Lehrer sind als KI-Pioniere an den Schulen unterwegs; ihr Wissen haben sie sich selbst angeeignet. Wir müssen aber allen Lehrern (…) vermitteln, wie sie ihren Unterricht mithilfe von Daten verbessern können und wie sie ihre pädagogischen Konzepte dadurch unterstützen können. (BT_244c, 2019, S. 14547)

Diese wiederkehrende Annahme, dass Investitionen in KI keiner weiteren Rechtfertigung bedürfen, verweist auf einen impliziten deterministischen Konsens, in dem KI als selbstverständlich notwendiger Bestandteil eines modernen Bildungssystems erscheint.

9.4.5.2. Mangelnde Rechtfertigung des Kompetenzerwerbs

Auch Forderungen des Wissenserwerbs über KI bleiben vielfach ohne explizite Rechtfertigung. Diese umfassen den Erwerb KI-bezogener Kompetenzen bei Lernenden, Lehrenden, Fachkräften sowie der gesamten Gesellschaft (GWK_1060, 2022, S. 105; LT-TH_497, 2018, S. 7; Wulf/CDU, LT-NI_664, 2019, S. 5799; LT-TH_836, 2021, S. 44 f.; LT-NW_895, 2021,

S. 164 f.; LT-SH_934, 2021, S. 25; WR_1175, 2022, S. 19). So fordert erneut die FDP, „dass Lehrpersonal bereits in der Ausbildung in ausreichendem Maße mit Methoden der Künstlichen Intelligenz vertraut und geschult wird“ (BT_288b, 2021, S. 7), und auch die GWK befasst sich mit einer „Stärkung der KI-Kompetenzen bei Studien- und Qualifizierungsangeboten“ (GWK_1052, 2020, S. 1). Dabei werden zwar handlungsleitende Ziele formuliert, jedoch ohne eine moralisch fundierte Rechtfertigung des Kompetenzbedarfs vorzulegen. Die rheinland-pfälzische Regierung hebt gar hervor, dass „es in frühen Stationen einer Bildungslaufbahn bereits einer spielerischen Annäherung an das algorithmische Denken“ (LT-RP_522, S. 2) bedürfe, um den Grundstein für spätere Programmierfähigkeiten zu legen. Projekte wie „KI Rollout“ (LT-RP_933, 2021, S. 21 f.) oder „Coding Detectives“ (LT-RP_522, S. 2) zeigen, dass entsprechende Ansätze mitunter bereits im Grundschulalter umgesetzt werden, wenngleich eine normative Rechtfertigung für die frühe Förderung von Programmierfähigkeiten ausbleibt. Analog setzt auch Niedersachsen auf eine „Förderung von Informatik [durch] die Fortbildungsreihe ‚Medienkompetenz an der Grundschule‘“ (LT-NI_798, 2021, S. 59) und sieht vor, Informatik „als Pflichtfach in allen Schulformen des Sekundarbereichs I“ (ebd., S. 60) zu verankern. Meyer (CDU) weitet die suggerierte Notwendigkeit des KI-Kompetenzerwerbs auf alle Bildungsstufen aus:

> Wir müssen dafür sorgen, dass gerade die junge Generation – es beginnt schon bei den Kindern im Kindergarten, geht über die Schulen und reicht letztlich bis hin zum lebenslangen Lernen – Hintergründe und Grundkenntnisse vermittelt bekommen. Es muss nicht gleich jeder Informatiker werden, aber man muss, um gesellschaftliche Entwicklungen zu verstehen, einen Zugang finden. (LT-SN_904, 2021, S. 2786)

Vor diesem Hintergrund werden die Einführung eines Pflichtfachs Informatik ebenso wie der Ausbau von KI-bezogenen Professuren, Forschungsschwerpunkten, Lehrveranstaltungen und Studiengängen als selbstverständlich präsentiert (BT_157b, 2019, S. 7; LT-HH_497, 2018, S. 2; LT-RP_522, S. 2; LT-NI_662, 2019, S. 4; Montag/FDP, LT-TH_736, 2020, S. 1517; LT-NI_798, 2021, S. 60 f.; WR_1127, 2019, S. 112; WR_1175, 2022, S. 38, 55), ohne dass ihre Notwendigkeit explizit gerechtfertigt wird. Diese Kompetenzen werden als elementar erachtet, wie Unions-Abgeordneter Steier betont: „das Problem ist ja nicht nur das Verstehen, sondern auch das Umsetzen einer neuen Technologie“ (BT_244c, 2019, S. 14548). Die Analyse zeigt dabei, dass diese Diagnosen und Forderungen häufig auf explizite, moralisch fundierte Rechtfertigungen verzichten. Stattdessen scheinen diese Entscheidungen und Maßnahmen durch unausgesprochene Prämissen geleitet zu werden: Erstens wird der KI-Einsatz einer deterministischen Annahme folgend als unumgänglich und zweitens der Erwerb entsprechender

Kompetenzen als selbstverständlich und wünschenswert vorausgesetzt. Die Debatte folgt damit an dieser Stelle einer Logik, in der die Einführung von KI keiner kritischen Auseinandersetzung über die Frage des „Ob“ mehr zu bedürfen scheint, sondern lediglich ein passender Rahmen durch entsprechende Bildungs- und Investitionsmaßnahmen geschaffen werden müsse.

Williamson (2016) kritisierte diese Forderung des „Learning to Code“ bereits als Ausdruck eines solutionistischen Denkens (siehe Abschnitt 6.4.4). Daran anknüpfend kann auf Basis der vorliegenden Erkenntnisse zwischen expliziten und impliziten Formen des Solutionismus differenziert werden. Während expliziter Solutionismus konkrete und auch diffuse Problemlösungsbestrebungen formuliert (siehe Abschnitte 9.4.2 und 9.4.3), entzieht sich der hier geäußerte implizite Solutionismus einer kritischen Rechtfertigung. Er scheint vielmehr auf der Annahme zu basieren, dass die erhobenen Forderungen weder kontingent noch rechtfertigungsbedürftig seien, sondern als selbstverständlich und unumstößlich gelten. Sie erscheinen als Ausdruck eines vermeintlichen Konsenses, wonach der KI-Einsatz per se positive Effekte nach sich ziehe. In diesen Segmenten fungiert der Solutionismus somit gewissermaßen als „Default-Option“ (Morozov, 2020), wodurch eine explizite Rechtfertigung nicht mehr erforderlich erscheint.

9.4.6. Effiziente Problemlöser: Der technikoptimistische Kompromiss von Industrie und Solution

Mit der Verbindung von Rechtfertigungen der Cité der Industrie und der Cité der Solution legt die Analyse einen weiteren zentralen Kompromiss der bildungspolitischen Debatte offen. Beide Cités eint eine ausgeprägt technikoptimistische Grundhaltung, auch wenn sie unterschiedlichen Logiken folgen: Die Cité der Industrie betont Innovation, Effizienz und wirtschaftliche Leistungsfähigkeit, die Cité der Solution fokussiert die Problemlösungsfähigkeit technologischer Systeme. Trotz dieser Differenz zielen beide auf einen produktiven und leistungsfähigen KI-Einsatz, sodass sich ihre Argumentationen häufig überlagern und wechselseitig stabilisieren.

Besonders deutlich tritt diese Verbindung in politischen Initiativen zutage, die KI als „Schlüsseltechnologie“ für Bildung und Wirtschaft gleichermaßen inszenieren (CDU, LT-HH_930, 2021, S. 1). Meyer (CDU) verbindet Effizienz und Problemlösung, wenn er betont: „Es geht um Anwendungen und Systeme, die Aufgaben lösen, welche bisher nur vom Menschen gelöst werden konnten und die jetzt durch Maschinenlernen (…) effizient ausgewertet werden können“ (LT-SN_904, 2021, S. 2586). Auch die Thüringer FDP-Fraktion verweist auf die systemprägende Bedeutung von KI, die „die Welt noch viel umfassender und schneller verändern [werde] als bisher (…) und ein zunehmend wichtiger Wirtschaftsfaktor“ (LT-TH_733, 2020, S.

6) sei. Diese Aussagen verdeutlichen, wie wirtschaftliche Modernisierung und gesellschaftliche Problembearbeitung in gemeinsamen Rechtfertigungen verschränkt werden.

Dies äußert sich ebenso bereits in der Präambel der Bund-Länder-Vereinbarung „Künstliche Intelligenz in der Hochschulbildung", in der KI als Technologie beschrieben wird, die „in fast allen Sektoren einsetzbar ist, erhebliche produktivitätserhöhende Effekte entfalten kann und damit eine entscheidende Wirkung auf die technologische Leistungsfähigkeit der Bundesrepublik Deutschland" (GWK_1052, 2020, S. 1) haben werde. Gleichzeitig formuliert die Förderinitiative zwei Ziele: „a) die Qualifizierung von zukünftigen akademischen Fachkräften durch die Implementierung von KI als Studieninhalt, b) die Verbesserung der Qualität, Leistungsfähigkeit und Wirksamkeit der Hochschulbildung durch den Einsatz von KI" (ebd.). Dabei verbinden sich industriepolitische und solutionistische Rechtfertigungslogiken zu einem gemeinsamen optimistisch-deterministischen Legitimationsrahmen: KI erscheint zugleich als Motor wirtschaftlicher Innovations- und Leistungsfähigkeit und als unumgängliches Mittel zur Modernisierung des Bildungswesens.

Auch in der hochschulpolitischen Praxis wird diese doppelte Rechtfertigungslogik deutlich. Der WR fordert, Hochschulen sollten innovative Technologien „nicht nur nutzen, sondern deren Entwicklung und Erprobung auch durch entsprechende Forschung unterstützen" (WR_1177, 2022, S. 94). Ziel sei ein institutioneller „Qualitätssprung", etwa durch Chatbots, virtuelle Labore oder Simultanübersetzungsprogramme. Somit streben solutionistische Nutzenversprechen und industrielle Innovationsförderung nach wechselseitiger Stabilisierung der Legitimation. Dies zeigt sich ebenso in der Rechtfertigung digitaler Prüfungsformate, die zugleich Effizienzversprechen einlösen und bildungspolitische Probleme juristischer Anfechtbarkeit analoger Prüfungen adressieren (HFD_1187, 2015, S. 44). Auch in der Unterrichtsgestaltung wird diese doppelte Stoßrichtung deutlich: Das Lernen solle durch den KI-Einsatz „erfolgreicher gestaltet" (HFD_1193, 2017, S. 7) und somit zugleich „einer veränderten Lernkultur und veränderten Lernformen Rechnung tragen" (KMK_1076, 2021, S. 14). Bildungspolitische Herausforderungen und KI-basierte Lösungen werden so eng gekoppelt, dass die Verbesserung der Bildungsqualität von der Einführung innovativer Systeme abhängig erscheint. Die FDP fordert entsprechend steuerliche Anreize für Unternehmen, die neue innovative Lernmethoden mit KI entwickeln (BT_288a, 2020, S. 4), und unterstreicht damit die gemeinsame Verantwortung von Industrie und Bildungspolitik (vgl. BT_113b, 2020, S. 184).

Daran anschließend äußert sich der solutionistisch-industrielle Kompromiss in der Verbindung gesellschaftlicher und wirtschaftlicher Nutzenargumentationen: „Von dem digital

transformierten Bildungssystem können alle profitieren: jede und jeder Einzelne durch bessere Beschäftigungsfähigkeit und Entwicklungschancen sowie Unternehmen durch eine verbesserte Standortqualität und eine hohe Innovationsfähigkeit“ (BT_113b, 2020, S. 303). Was individuell nützlich sei, solle zugleich wirtschaftlicher Innovation dienen, wie auch in datenpolitischen Positionierungen zum Ausdruck kommt. So fordert Schön (Union) „einen innovationsoffenen Umgang mit Daten, der die Persönlichkeitsrechte der Menschen schützt, aber gleichzeitig unserem wirtschaftlichen und gesellschaftlichen Bedürfnis nach Nutzung von künstlicher Intelligenz gerecht wird“ (BT_113d, 2020, S. 23758). Hier verbinden sich solutionistische Problemlösungserwartungen, industrielle Effizienzlogiken und staatsbürgerliche Schutzargumente zu einer gemeinsamen Legitimationsformel. An die Rückbindung technologischer Innovation an gesellschaftliche Ziele knüpft auch SPD-Abgeordnete Friedel an:

> Die Technologie ist vorhanden; wir müssen jetzt irgendwie dabei sein. Was mir ein wenig fehlt, ist der Blick auf das Potenzial, was künstliche Intelligenz ausmacht. (…) Wozu machen wir das Ganze eigentlich? Eine der wichtigen Antworten ist, die Arbeitszeit- und Produktivitätsgewinne in neue Lebensqualität zu stecken. (LT-SN_904, 2021, S. 2789)

Die Unumgänglichkeit des KI-Einsatzes wird so durch Effizienz und die Verbesserung des gesellschaftlichen Status quo gleichermaßen gerahmt. Insgesamt zeigt sich damit, dass dieser Kompromiss von einer hohen Anschlussfähigkeit getragen wird. Solutionistische Problembearbeitungen benötigen industrielle Umsetzungs- und Verwertungskompetenz, während die Industrie ihre Innovationsbestrebungen zunehmend durch gesellschaftlich anerkannte Nutzenversprechen legitimiert. Beide Cités verfolgen unterschiedliche Rechtfertigungsansätze, arbeiten aber an einem gemeinsamen Zukunftsbild: Technologiegestützter Fortschritt erscheint zugleich als ökonomisch rational und gesellschaftlich wünschenswert.

9.4.7. Wertebasierte Problemlösung: Ein instrumenteller Kompromiss von Staatsbürger und Solution

Mit dem Kompromiss zwischen der Cité des Staatsbürgers und der Solution identifiziert die Analyse einen letzten zentralen Kompromiss der bildungspolitischen Debatte. Ob es sich dabei jedoch um einen stabilen Kompromiss im Sinne gegenseitiger Anerkennung oder eher um eine strategische Instrumentalisierung staatsbürgerlicher Prinzipien zugunsten solutionistischer Interessen handelt, gilt es an späterer Stelle kritisch zu diskutieren (siehe Abschnitt 10.2.4). So zeigt die Verflechtung der Cités, dass politische Akteur:innen solutionistischen Rechtfertigungen KI-basierter Problemlösung durch die Inklusion staatsbürgerlicher Werte wie Chancengleichheit oder Teilhabe zusätzliche politische Durchsetzungskraft verleihen wollen. Der Kompromiss folgt dabei dem Muster, dass es staatsbürgerliche Ansprüche zu erfüllen gelte, damit

dem KI-Einsatz nichts mehr im Wege stehe (BT_113b, 2020, S. 100; Gerber/Die Grünen, LT-SN_904, 2021, S. 2791 f.). Suding (FDP) fordert entsprechend in Bezug auf Learning Analytics, dass ein klarer Rechtsrahmen notwendig sei, um ihr volles Potenzial entfalten zu können (BT_244c, 2019, S. 14547). Völlers (SPD) schließt daran an: „Wenn all diese [ethischen und datenschutzrechtlichen] Fragen geklärt sind, dann wird Learning Analytics kommen und mit all seinen Vorzügen unsere Schulen begeistern" (BT_244c, 2019, S. 14627). Staatsbürgerliche Werte dienen hier weniger der kritischen Prüfung als vielmehr der Absicherung eines präferierten technologischen Lösungsansatzes und werden so gewissermaßen zum Legitimitätsvehikel des solutionistischen Fortschrittsdenkens.

Diesem argumentativen Muster bedient sich auch der Antrag „Alphabetisierung durch Künstliche Intelligenz" der FDP, der die unzureichende Chancengleichheit im Bildungswesen als Problem benennt (siehe Abschnitt 9.5.1) und KI als Lösung präsentiert, indem diese zur Literarisierung und Grundbildung beitragen solle (BT_288b, 2021, S. 2 f.). Alternative Maßnahmen werden jedoch nicht diskutiert, stattdessen wird Learning Analytics als selbstverständlicher Bestandteil der Antwort präsentiert und damit die Cité des Staatsbürgers funktional in den Dienst der KI-Lösung gestellt. Dieses Muster zeigt sich ebenso im Kontext der Vision einer nationalen Bildungscloud. Eine solche Infrastruktur soll als Grundlage einer Vielzahl von KI-Lösungen wie etwa intelligenten Tutorensystemen dienen. Informatiker Meinel betont, dass diese humboldtsche Bildungsideale einlösen könne, während analoge Lernformen sogar Chancenungleichheit verstärken würden: „Erst mit Hilfe neuester Digitaltechnologien können die humboldtschen Erwartungen (…) realisiert werden" (HFD_1193, 2017, S. 8). Damit wird das humboldtsche Ideal zur Begründung für ein solutionistisch geprägtes Infrastrukturprojekt herangezogen. In diese Logik fügt sich der MINT-Aktionsplan 2.0 ein, vor dessen Hintergrund die FDP argumentiert: „digitale Bildung ermöglicht beste Bildungs- und Teilhabechancen für alle" (Promny/FDP, LT-HE_1011, 2022, S. 8583) und daher brauche das Land „MINT-Begeisterte, um die Herausforderungen der Zukunft meistern zu können" (ebd.). Die Technologieförderung erscheint damit als Voraussetzung gesellschaftlicher Gleichheit, während nicht-digitale Optionen nicht erwogen werden.

Die Einbettung solutionistischer Maßnahmen in staatsbürgerliche Zielsetzungen zeigt sich auch in der Debatte über KI-Kompetenzen, die als Befähigung zur demokratischen Teilhabe gerechtfertigt werden. Der Enquete-Bericht betont: „Nur wer mitentwickelt, kann auch effektiv mitbestimmen" (BT_113b, 2020, S. 336). Demnach wird das Erlernen technischer Kompetenzen zur Voraussetzung politischer Mitsprache erklärt – ein Anspruch, der tief in staatsbürgerlicher Rhetorik verankert ist, tatsächlich jedoch einem solutionistischen Imperativ zu folgen scheint. Auch

der Abbau geschlechtsspezifischer Ungleichheiten wird durch KI-Angebote adressiert, etwa durch gendersensible Lernformate (BT_113b, 2020, S. 340 f.). Zwar werden staatsbürgerliche Werte affirmativ in solutionistische Rechtfertigungen eingebunden, doch unterliegt ihre Verwendung häufig scheinbar einer solutionistischen Lösungseuphorie.

Forderungen nach einem verbindlichen Rechtsrahmen und dem Schutz der Persönlichkeitsrechte existieren zwar in diesem Kontext (siehe Abschnitt 9.5.1.3), dienen aber vielfach eher der Ermöglichung technischer Innovationen als ihrer Regulierung. So formuliert Friedel (SPD): „Ich denke, dass die bestehenden Risiken mit einer vernünftigen und klaren Regulierung recht gut in den Griff zu bekommen sind. (…) Künstliche Intelligenz kann, wie jede neue Technologie, das Leben der Menschen erleichtern“ (LT-SN_904, 2021, S. 2789). Auch hier dominiert die Vision des Nutzens, während die Cité des Staatsbürgers vor allem der Stabilisierung der Legitimation dient, statt der kritischen Begrenzung, was insbesondere in der fehlenden Definition der konkreten Risiken deutlich wird.

Auf dieser Grundlage lässt sich ein ambivalentes Bild zeichnen: Einerseits findet eine klare Bezugnahme auf staatsbürgerliche Werte wie Chancengleichheit, Teilhabe und Legalität statt. Andererseits bleibt unklar, ob diese Werte tatsächlich handlungsleitend sind oder lediglich rhetorisch eingebunden werden, um solutionistische Maßnahmen durchzusetzen. Soziologin Hartong warnt explizit davor, dass solutionistische Akteur:innen gezielt staatsbürgerliche Rechtfertigungen mobilisieren würden, um ihre Positionen durchzusetzen und sich gegen Kritik abzusichern:

> Während solche Akteure erfolgreich mit ihrer Neutralität werben, (…) lassen [sie] aber gleichzeitig keinen Zweifel daran, dass digitale Produkte unbedingt genutzt werden müssen, mehr Markt geschaffen werden muss – und platzieren diese Idee zunehmend erfolgreich in Politik und Schulen. Dabei haben sie es inzwischen quasi perfektioniert, mit Partizipation, demokratischer Teilhabe und Gemeinschaft zu argumentieren, auch um potentieller Kritik von Anfang an den Wind aus den Segeln zu nehmen (…). (GEW_1244, 2019, S. 16)

Hartong stellt somit die Instrumentalisierung der staatsbürgerlichen Cité bloß, welche in dieser Logik nicht als Ziel, sondern vielmehr als Mittel zur Durchsetzung digitaler Lösungen eingespannt wird. Vor diesem Hintergrund scheint die Frage berechtigt, ob an dieser Stelle tatsächlich ein stabiler Kompromiss im Sinne gegenseitiger Anerkennung zweier Cités vorliegt oder ob die Cité des Staatsbürgers lediglich als strategische Ressource im Dienste solutionistischer Problemlösungsrhetorik dient.

9.5. Kritik und Konfliktlinien der moralischen Rechtfertigung des Einsatzes Künstlicher Intelligenz

Nachdem die Analyse bis zu diesem Punkt transparent gemacht hat, auf welcher moralischen Grundlage der KI-Einsatz in der deutschen bildungspolitischen Debatte gerechtfertigt wird, richtet sich der Fokus nun auf die angeführte Kritik sowie auf die Konfliktlinien dieser Auseinandersetzung (FF3). Während die Rechtfertigungen unterschiedliche normative Orientierungen sichtbar machen, verweisen die kritischen Stimmen auf grundlegende Spannungen, die sich aus der politischen, pädagogischen und gesellschaftlichen Einbettung von KI ergeben. Besonders deutlich wird dabei, dass sich die Kritik am häufigsten aus der Perspektive der Cité des Staatsbürgers und der Cité des Hauses artikuliert, wie auch in Anbetracht der Valenz der Cités deutlich wird (siehe Abbildung 5). Beide eröffnen zentrale Gegenpositionen, indem sie nicht nur den KI-Einsatz selbst in Frage stellen, sondern zugleich die Rechtfertigungsmuster anderer Cités problematisieren, wie im Rahmen von Existenzprüfungen zum Ausdruck kommt (siehe Abbildung 6). Im Folgenden werden die wesentlichen Kritikpunkte aus Sicht dieser beiden Cités sowie die daraus hervorgehenden Konfliktlinien zwischen unterschiedlichen Rechtfertigungsnarrativen als existenzielle Prüfungen rekonstruiert, um die Tiefe der Auseinandersetzung um Legitimität und Grenzen des KI-Einsatzes im Bildungssystem zu analysieren.

9.5.1. Staatsbürgerliche Kritik der Verletzung von Grundrechten und demokratischer Werte durch Künstliche Intelligenz

Die staatsbürgerliche Cité bildet neben der Cité des Hauses eine zentrale moralische Grundlage für die Kritik des KI-Einsatzes im Bildungswesen. Ihr Schwerpunkt liegt auf der Wahrung demokratischer Werte, individueller Rechte und gesellschaftlicher Gleichheit. Die Analyse zeigt, dass sich die staatsbürgerliche Kritik sowohl auf die Funktionsweise von KI-Anwendungen bezieht als auch auf die Rechtfertigungsansprüche anderer Cités, die aus dieser Perspektive als unvereinbar mit staatsbürgerlichen Prinzipien gelten. Die Ergebnisse verdeutlichen, dass die Kritik in zwei Formen auftritt. Einerseits führt sie zu Realitätsprüfungen, die im Sinne reformistischer Kritik auf Veränderungen innerhalb bestehender Strukturen zielen. Sie fordern etwa eine verlässliche Regulierung und transparente Verfahren, um einen diskriminierungsfreien Einsatz von KI sicherzustellen. Diese Kritik richtet sich auf die Umsetzung bestehender staatsbürgerlicher Normen und die Anpassung zwischen gesellschaftlichem Anspruch und politischer Praxis. Andererseits artikuliert sich eine radikale Kritik, die die Legitimität des KI-Einsatzes grundsätzlich infrage stellt und zu Existenzprüfungen führt. Sie verweist etwa auf die Unvereinbarkeit algorithmischer Bewertung mit humanistischen Bildungsidealen und lehnt damit zugleich die Rechtfertigungen anderer Cités, insbesondere der Solution, der Industrie und des Marktes, ab.

Im Folgenden werden die zentralen Kritikpunkte der staatsbürgerlichen Cité im bildungspolitischen Diskurs rekonstruiert. Sie reichen von der Gefährdung des Gerechtigkeitsanspruchs durch diskriminierende KI-Anwendungen über die Bedrohung humanistischer Bildungsideale bis hin zur Verletzung von Persönlichkeitsrechten und der Kritik an Überwachungs- und Kontrolllogiken. Existenzprüfungen und Konflikte mit anderen Cités entstehen vor allem dort, wo technokratische Steuerungsabsichten, solutionistische Problemlösungslogiken und industrielle Effizienzorientierungen das Verständnis von Bildung als Raum kritischer Selbstentfaltung herausfordern. Abschließend wird die Realitätsprüfung der Cité des Staatsbürgers im Kontext der Unsicherheit über Technikfolgen und Regulierungsfragen dargestellt.

9.5.1.1. Die Gefährdung des Gerechtigkeitsanspruchs durch mangelnde Objektivität und drohende Diskriminierung

Die Wahrung demokratischer Werte bildet den zentralen Maßstab der staatsbürgerlichen Cité. Entsprechend richtet sich ihre Kritik am KI-Einsatz auf die Gefahr, dass automatisierte Entscheidungsprozesse Grundprinzipien der Gleichheit und Gerechtigkeit untergraben. Im Bildungssektor steht dabei die Frage der Chancen- und Bildungsgerechtigkeit im Mittelpunkt. So zeigt die Analyse, dass in der bildungspolitischen Debatte die sozialen und demokratischen Folgen von KI-Anwendungen hinterfragt werden. Wiederkehrend ist die Befürchtung, dass algorithmische Systeme bestehende Ungleichheiten reproduzieren oder gar verstärken könnten (GEW_1239, 2019, S. 1 ff.; GEW_1244, 2019, S. 13; LT-BW_877, 2021, S. 43; Gerber/Die Grünen, LT-SN_904, 2021, S. 2789; Kula/Die Linke, LT-HE_1011, 2022, S. 8591; TAB_1327, 2021, S. 3; TAB_1324, 2022, S. 3). Diese Diskriminierungsgefahr wird sowohl im Hinblick auf Geschlechterdiversität (BT_113b, 2020, S. 350; BT_162a, 2018, S. 8; BT_157a, 2018, S. 4; LT-SH_934, 2021, S. 18) als auch auf sozioökonomische Unterschiede im Zugang zu Bildung und in der Prognosegenauigkeit betont (GEW_1239, 2019, S. 1 ff.; GEW_1244, 2019, S. 12 f.; GEW_1268, 2020, S. 1 ff.). Die Gefahr der bewussten Diskriminierung durch Manipulation wird hingegen nur vereinzelt thematisiert (Domscheit-Berg/Die Linke, BT_157c, 2018, S. 8101; GEW_1338, 2021, S. 28; WR_1177, 2022, S. 31 f.).

So warnt die Bundesregierung: „Gesellschaftliche Ungleichheit und Diskriminierung von Individuen oder Personengruppen darf nicht fortgeschrieben oder gar verstärkt werden, indem in die Programmierung oder durch die Trainingsdaten tendenziöse Vorfestlegungen oder Diskriminierungen einfließen“ (BT_162b, 2019, S. 16). Fragen der Fairness werden ebenso von der Enquete-Kommission in den Vordergrund gerückt, die im Zuge einer reformistischen Kritik fordert, dass KI-gestützte Entscheidungen regelmäßig auf ihre Diskriminierungsfreiheit hin überprüft werden müssen (BT_113b, 2020, S. 41). Neben der Einhaltung gesetzlicher Vorgaben

sollen Diversität und die Beteiligung der Betroffenen, insbesondere der Lernenden, berücksichtigt werden. Zudem sollen Individuen befähigt werden, sich gegen diskriminierende Praktiken zur Wehr zu setzen, um ein faires und inklusives System zu sichern (BT_113b, 2020, S. 63).

Expert:innengremien kritisieren in diesem Zusammenhang jedoch nicht nur das Diskriminierungspotenzial, sondern verweisen auch auf die grundsätzliche Fehlannahme neutraler und objektiver KI-basierter Entscheidungsfindung (TAB_1327, 2021, S. 5). Dies kann sowohl in der Funktionalität der Systeme als auch in den Input-Daten begründet liegen, wobei sowohl die Messbarkeit als auch die Auswahl der Daten kritisiert werden (Brünler/Die Linke, LT-SN_904, 2021, S. 2788). So gilt es unter anderem, zu berücksichtigen, dass „eine Lernplattform (..) nicht alle Parameter einbeziehen [kann], die zu einer gerechten oder objektiven Bewertung führen“ (HFD_1210, 2021, S. 2 f.).

Der Konflikt um Chancen- und Bildungsgerechtigkeit zwischen staatsbürgerlicher Sorge und solutionistischer Rechtfertigung

Anknüpfend an das Diskriminierungspotenzial und die Infragestellung der Objektivität von KI betrifft ein zentraler Konfliktpunkt in der bildungspolitischen Debatte um KI die Frage, ob ihr Einsatz zur Sicherung oder zur Gefährdung von Chancen- und Bildungsgerechtigkeit beiträgt (GEW_1338, 2021, S. 5; HFD_1210, 2021, S. 1 ff.). Damit verdichtet sich eine Existenzprüfung zwischen staatsbürgerlichen Bedenken hinsichtlich sozialer Ungleichheit und solutionistischen Rechtfertigungen, die KI als Instrument zur Verbesserung des Bildungswesens darstellen. Die Existenzprüfung wird dabei durch die radikale Kritik sichtbar, die die Legitimität des jeweils anderen moralischen Rechtfertigungsanspruchs und damit die bestehende Ordnung infrage stellt (siehe Abschnitt 5.2.2; Boltanski, 2010).

Auf der einen Seite wird dabei das Potenzial digitaler und KI-gestützter Verfahren betont, etwa durch Learning Analytics oder automatisierte Prüfungsprozesse, individuelle Förderung zu ermöglichen und damit Bildungsgerechtigkeit zu steigern (HFD_1191, 2016, S. 36; LT-HE_1011, 2022, S. 8583 f.; TAB_1333, 2021, S. 6; WR_1177, 2022, S. 95). In diesem Verständnis gilt die KI-basierte Datennutzung als Voraussetzung einer objektiveren Leistungsbewertung und einer gerechteren Ressourcenverteilung. So formuliert die FDP programmatisch: „nur durch eine datengestützte Schulentwicklung können wir Chancengerechtigkeit in diesem Land auch tatsächlich voranbringen. (…) Wir wollen zum Meister der Chancengerechtigkeit werden und zugleich ein hohes Leistungsniveau erreichen“ (Promny/FDP, LT-HE_1011, 2022, S. 8584).

Auf der anderen Seite artikuliert sich in der staatsbürgerlichen Kritik die Sorge, dass eine technikzentrierte Bildungspolitik normative Fragen von Gleichheit und Inklusion vernachlässigt. Vertreter:innen der Grünen kritisieren etwa eine innovationspolitische Priorisierung der FDP nach dem Motto „KI first, Digitalisierung second, Standards third, Bedenken fourth or fifth or whatever" (Stumpp/Die Grünen, BT_244c, 2019, S. 14551). Diese sei grundsätzlich fehlerhaft und „im Hinblick auf Bildungs- und Zukunftschancen von Kindern geradezu ein Frevel" (ebd.). Während einerseits ein Mangel an wissenschaftlichen Erkenntnissen darüber festgestellt wird, ob der KI-Einsatz tatsächlich zu besserem Lernerfolg oder geringerer Bildungsungleichheit führt (ebd.; WR_1177, 2022, S. 38), lehnt die GEW die Annahme der Verringerung der Bildungsungleichheit durch den Einsatz von KI gänzlich ab (GEW_1244, 2019, S. 13). Kritisiert wird, dass eine Vereinheitlichung von Lernprozessen und deren algorithmische Bewertung bestehende Ungleichheit und Selektivität eher verstärken könnte (GEW_1239, 2019, S. 2; GEW_1268, 2020, S. 2). Die Enquete-Kommission warnt im Zuge dessen vor einer frühen Determinierung von Lebensverläufen, die einen unethischen Eingriff in Bildungsbiografien darstelle. Insbesondere die KI-basierte Studienabbruchsprognose wird als Beispiel dafür genannt, wie algorithmische Eingriffe unbeabsichtigte soziale Folgen haben können, wie zum Beispiel Demotivation oder Stigmatisierung junger Menschen (BT_113b, 2020, S. 339).

Die staatsbürgerliche Kritik richtet sich somit gegen die Objektivitäts- und Neutralitätsannahme algorithmischer Entscheidungen (Stumpp/Die Grünen, BT_244c, 2019, S. 14551; Kula/Die Linke, LT-HE_1011, 2022, S. 8591). Vielmehr wird auf die soziale Konstruiertheit der Daten hingewiesen, die durch ihre Erhebungslogik Machtverhältnisse reproduzieren. Die GEW betont,

> dass Daten niemals ein neutrales, entpolitisiertes, technologisch verdichtetes Abbild von Realität sind (als was sie nämlich gerade bei Lerntechnologien oftmals verkauft werden), sondern zum einen immer von der Umgebung, in der sie entstehen, massiv beeinflusst sind, dass Daten zum anderen aber auch immer Wirklichkeit erschaffen und damit eine oftmals unterschätzte Macht ausüben. (GEW_1244, 2019, S. 7)

In dieser Perspektive ist die solutionistische Annahme einer technikvermittelten Gerechtigkeit selbst Teil eines politischen Problems, da durch den KI-Einsatz im Bildungswesen sowohl eine Fortschreibung historisch gewachsener Diskriminierungsmuster drohe (LT-BW_877, 2021, S. 43; Brünler/Die Linke, LT-SN_904, 2021, S. 2787 f.), als auch neue Ungleichheiten durch die Logik der Datenmodellierung erzeugt würden: „Ungleichheit muss vielmehr produziert

werden, um Daten algorithmisch unterschieden [sic!], sortieren, gruppieren, hierarchisieren und daraus Entscheidungen ableiten zu können“ (GEW_1244, 2019, S. 13).

Neben diesen strukturellen Effekten der Datenproduktion können ebenso Disparitäten im Zugang zu Technologien sowie Nutzungskompetenzen soziale Ungleichheit verstärken. Liegen diese Kompetenzen nicht vor, „wird der Zugang zu zukunftsträchtigen Berufen bzw. Einkommensmöglichkeiten erschwert. Klar ist, dass der Zugang zu digitalen Technologien und die Fähigkeit, sie zu nutzen, wichtige Bedingungen für gesellschaftliche und berufliche Teilhabe sind“ (TAB_1334, 2021, S. 7). Durch diese Feststellung des TAB werden neue, technikvermittelte Formen der Ungleichheit sichtbar, die sich nicht durch weitere Digitalisierung lösen lassen, sondern politische Gegensteuerung erfordern.

Insgesamt zeigt sich damit ein Spannungsfeld zwischen zwei konkurrierenden Deutungen von Bildungsgerechtigkeit: Während solutionistische Rechtfertigungen Gerechtigkeit als technisches Steuerungsziel begreifen, das durch datenbasierte Verfahren erreichbar ist, verstehen staatsbürgerlich argumentierende Akteur:innen sie als normatives Prinzip, das losgelöst von technologischen Lösungen zu verfolgen ist, wie der Grünen-Abgeordnete May in einer Debatte des hessischen Landtags betont: „Für uns (…) steht im Vordergrund, dass wir die Bildungsgerechtigkeit und die Individualisierung von Unterricht weiterentwickeln wollen, um möglichst viel Bildungserfolg für alle Schülerinnen und Schüler zu ermöglichen. Diesem Ziel muss sich auch die Digitalisierung beugen“ (LT-HE_1011, 2022, S. 8593). Damit markiert die KI-Debatte um Bildungsgerechtigkeit nicht nur eine Auseinandersetzung über technische Möglichkeiten, sondern eine Existenzprüfung unterschiedlicher Gemeinwohlvorstellungen solutionistischer und staatsbürgerlicher Ansprüche: ob Gerechtigkeit durch Technologie erzeugt werden kann oder ob sie der Technologie selbst als Grenze dienen muss.

9.5.1.2. Die Realitätsprüfung des humanistischen Bildungsverständnisses im Kontext des Einsatzes Künstlicher Intelligenz

Das gesellschaftliche Bildungsverständnis wird durch den KI-Einsatz zunehmend herausgefordert. Im Zentrum der Cité des Staatsbürgers steht die Anerkennung und Bewährung eines humanistischen Bildungsideals, das Persönlichkeitsentwicklung, Mündigkeit, kritisches Denken, Kreativität und Souveränität hervorhebt (HFD_1202, 2021, S. 54; HFD_1206, 2021, S. 2). So markiert dieses Bildungsverständnis in der KI-Debatte sowohl eine Realitätsprüfung der Cité des Staatsbürgers als auch eine Existenzprüfung dieser staatsbürgerlichen Ansprüche gegenüber den Cités der Solution und der Industrie, welche die Vermittlung technischer Kompetenzen beziehungsweise ein effizienzorientiertes Bildungsverständnis zugrunde legen. Zunächst

wird jedoch die Realitätsprüfung des staatsbürgerlichen Bildungsverständnisses dargelegt. Dabei wird, um dies noch einmal in Erinnerung zu rufen, die Legitimität der Ansprüche der Cité des Staatsbürgers grundsätzlich anerkannt, ihre konkrete Ausgestaltung ist jedoch Gegenstand kontinuierlicher reformistischer Kritik und Anpassung, wie es die pragmatische Soziologie für solche Bewährungsproben beschreibt (Boltanski & Thévenot, 1991/2018; Smith Ochoa et al., 2021).

Insbesondere die Fähigkeit zum kritischen Denken steht dabei im Fokus der Realitätsprüfung der Cité des Staatsbürgers. Während Enquete-Kommission und TAB betonen, KI und mobile Lernformen könnten „kritisches Denken bei Lehrenden sowie Lernenden fördern" (TAB_1335, 2020, S. 81; s. a. BT_113b, 2020, S. 350) und damit staatsbürgerliche Werte stützen, verweisen insbesondere gewerkschaftliche Akteur:innen darauf, dass gerade die Verbreitung leistungsfähiger KI-Systeme eine gezielte Förderung analytischer und kritischer Reflexionskompetenzen erforderlich macht (GEW_1244, 2019, S. 14). Vor dem Hintergrund dieser als irreversibel wahrgenommenen Entwicklung hebt auch das HFD die Ausbildung analytischer und kritischer Reflexionsfähigkeiten als Schlüsselkompetenz für Lehrende und Lernende hervor, um KI-Anwendungen bewerten und steuern zu können (HFD_1317, 2022, S. 5). Daraus folgt eine Verschiebung im staatsbürgerlichen Rechtfertigungsmodus: Der Verweis auf das humanistische Bildungsideal legitimiert nicht den KI-Einsatz an sich, sondern begründet eine Priorität auf kritische Urteilsbildung als Schutz- und Orientierungsressource angesichts unvermeidlicher Technikintegration und weitreichender Folgen (HFD_1202, 2021, S. 50). Auch im parlamentarischen Diskurs spiegelt sich dieses Ringen um die Bedeutung des kritischen Denkens sowie eines reflektierten Umgangs mit digitalen Medien wider (Jongen/AfD, BT_113c, 2018, S. 4157; Felser/AfD, BT_113d, 2020, S. 23757; Jankowski/AfD, LT-TH_850, 2021, S. 3546).

Die Realitätsprüfung zeigt somit, dass die Forderung nach kritischem Denken in der staatsbürgerlichen Rechtfertigung zwar als weitgehender Konsens in Bezug auf Bildungsziele der Cité des Staatsbürgers gilt, zugleich aber umstritten bleibt, ob und inwiefern der Einsatz von KI selbst zu dessen Förderung beitragen kann oder vielmehr zusätzliche Herausforderungen hervorruft.

Der Konflikt zwischen solutionistischer Kompetenzvermittlung, falscher Problempriorisierung und dem staatsbürgerlichen Bildungsideal

Die Debatte verschärft sich zur Existenzprüfung dort, wo die Cité der Solution den Vorrang technischer Problemlösungen und den Vorrang technischer Kompetenzvermittlung reklamiert. Kritisiert wird, dass die zunehmende Leistungsfähigkeit von KI dazu geführt habe, dass ihr

Einsatz häufig nicht mehr sei als ein oberflächliches „Anstupsen“ im Unterricht, welches nicht ausreichend Raum für echtes, kritisches Denken biete (GEW_1244, 2019, S. 14). Enquete-Kommission und GEW warnen daher, dass solche Systeme eine passive Lernhaltung fördern können, die Konformität nahelegt und eine persönliche Entfaltung sowie kritische Auseinandersetzung mit gesellschaftlichen Entwicklungen erschwert (BT_113b, 2020, S. 339; GEW_1265, 2020, S. 2). Gleichzeitig wird „das Fehlen diskursiver und kommunikativer Auseinandersetzungen mit dem Unterrichtsgegenstand“ (Brünler/Die Linke, LT-SN_904, 2021, S. 2788) durch einen KI-Einsatz moniert. Das HFD akzentuiert dieses Spannungsfeld zwischen der gewünschten Anregung zu mündigen Entscheidungen und der Gefahr einer „Trübung der individuellen Urteilskraft“ (HFD_1202, 2021, S. 54) durch unreflektierte KI-Nutzung.

In diesem Kontext verschiebt sich die Auseinandersetzung um das Bildungsideal zur Debatte über Kompetenzen. Solutionistische Positionen rechtfertigen die Ausbildung informatischer Inhalte und die Einführung eines Pflichtfachs Informatik als unverzichtbar für moderne Bildung (siehe Abschnitt 9.4.5.2). Demgegenüber betonen staatsbürgerlich argumentierende Akteur:innen ein fächerübergreifendes Konzept der Medienbildung, das pädagogische, analytische, strukturelle und informationsbezogene Fähigkeiten mit kritischer Reflexion und partizipativem Handeln verbindet (GEW_1338, 2021, S. 18). Die GEW lehnt ein Pflichtfach Informatik ab und verteidigt ein umfassendes Medienbildungskonzept „gegen die technokratischen Auffassungen ‚digitaler Bildung‘ und Konzepten wie Learning Analytics“ (GEW_1338, 2021, S. 16). Die Auseinandersetzung ist damit nicht nur curricular, sondern normativ: Es geht um die Definition der „richtigen“ Kompetenzen in digitalen Gesellschaften (BT_113b, 2020, S. 341; TAB_1335, 2020, S. 28 f.; GEW_1244, 2019, S. 22; Raulfs/SPD, LT-NI_547, 2019, S. 3913) und um die Frage, ob Technik Gestaltungsmittel oder Gestaltungsziel von Bildung sein soll.

Die staatsbürgerliche Kritik richtet sich darüber hinaus gegen die falsche Priorisierung von Problemen. Dabei wird die bereits in der solutionistischen Bewährungsprobe zum Ausdruck gekommene Relativierung der KI als Werkzeug aufgegriffen (siehe Abschnitt 9.4.4). KI könne lediglich als unterstützendes Hilfsmittel einen Beitrag beispielsweise zur Inklusion und zur Überwindung gesellschaftlicher Herausforderungen leisten, jedoch nicht als Allheilmittel fungieren, etwa beim Abbau von Vorurteilen gegenüber Menschen mit Behinderungen. Die grundlegenden gesellschaftlichen Gestaltungsaufgaben blieben somit weiterhin relevant (LT-NW_753, 2020, S. 143). Ferner wird kritisiert, dass bestehende Probleme wie der Lehrkräftemangel oder die Heterogenität der Lerngruppen durch Learning Analytics nicht adäquat gelöst werden können, sondern vielmehr staatsbürgerliche Bildungsideale, insbesondere die Förderung einer kritisch denkenden Ausbildung, unterminiert würden (GEW_1272, 2021, S. 2). Auch

die Auseinandersetzung mit KI-basierten Weiterbildungsplattformen wie „Milla“ verdeutlicht, dass die seit Jahrzehnten bestehenden Probleme in der Weiterbildung, etwa prekäre Beschäftigungsverhältnisse und unzureichende finanzielle Ressourcen, aus Sicht der GEW nicht durch den Einsatz von KI kompensiert werden können (GEW_1250, 2019, S. 1). Grünen-Abgeordnete Stumpp fasst daher mit Verweis auf das Problem sozialer Ungleichheit für den Lernerfolg zusammen: „wir haben in unserem Bildungssystem doch ganz andere und deutlich wichtigere Herausforderungen als einen unzureichenden Einsatz von KI“ (Stumpp/Die Grünen, BT_244c, 2019, S. 14551). Und auch Unions-Abgeordneter Tiemann hebt mit Blick auf solutionistische Problemlösungsversprechen der FDP durch den Einsatz von Learning Analytics eine fehlerhafte Problemdefinition im Bildungskontext hervor, die „die Realitäten und Herausforderungen an unseren Schulen verkenne(…)“ (Tiemann/Union, BT_244c, 2019, S. 14626).

Aus staatsbürgerlicher Sicht verkürzt der solutionistische Bildungsansatz den Bildungsbegriff, indem er technische Kompetenzvermittlung über die Förderung von Mündigkeit und kritischem Denken stellt. Bildung verliert damit ihre Bedeutung als Raum demokratischer Selbstbestimmung. Die Existenzprüfung zeigt sich somit in zweifacher Hinsicht: als Konflikt um das Verständnis von Bildung und als Kritik an der politischen Priorisierung technischer Innovation gegenüber sozialen Gestaltungsaufgaben. In beiden Fällen verteidigt die Cité des Staatsbürgers das humanistische Bildungsideal gegen die Reduktion von Bildung auf technologische Lösbarkeit.

Der Konflikt zwischen industrieller Effizienzorientierung und Leistungsmetrisierung und staatsbürgerlicher Fürsorge im Bildungswesen

Parallel dazu steht das staatsbürgerlich gerechtfertigte humanistische Bildungsideal in einem grundlegenden Spannungsverhältnis zur Cité der Industrie, die Bildung als effizienten, outputorientierten Prozess versteht und zu diesem Zweck eine Metrisierung und Quantifizierung von Lernprozessen verfolgt. In den untersuchten Dokumenten wird deutlich, dass sich dieser Konflikt um die Frage dreht, ob Effizienz- und Leistungssteigerung überhaupt legitime Bildungsziele darstellen können, wenn sie die Autonomie und Selbstbildung der Lernenden gefährden. In der parlamentarischen Debatte über den FDP-Antrag „Smart Germany – Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen“ wird diese Frontstellung besonders deutlich:

> Bildung bedeutet im humanistischen Sinne, (…) eine reflektierende, abwägende Auseinandersetzung mit sich selbst und der Welt, die im besten Falle dazu führt, seinen Platz in dieser Welt zu finden. Bildung bedeutet nicht (…): Wie bekommt man mit möglichst

> wenig Ressourcen möglichst viel Stoff möglichst schnell in die Köpfe? Bildung in einer digitalen Gesellschaft muss ein sozialer Prozess kollaborativen Lernens und Arbeitens sein, der eine kritische Sicht auf die Dinge ermöglicht und junge Leute in die Lage versetzt, kritische Distanz zu wahren. (Bull-Bischoff/Die Linke, BT_244c, 2019, S. 14627)

Auch jenseits parteipolitischer Linien wird dieser staatsbürgerliche Anspruch betont: So verweist Espendiller von der AfD-Fraktion auf „das Bildungsideal von Humboldt: Schüler sollen mündige Erwachsene werden, die ganzheitlich gebildet sind“ (BT_244c, 2019, S. 14549). Und auch Abgeordnete Stumpp von den Grünen kritisiert eine Verkürzung des Bildungsbegriffs durch die Fokussierung auf KI-gestützte Effizienzsteigerung, die die Persönlichkeitsentwicklung als zentrales Bildungsziel vernachlässige (BT_244c, 2019, S. 14551).

Diese Kritik richtet sich gegen eine Bildungspolitik, die zunehmend als Vehikel wirtschaftlicher Standortförderung verstanden wird. Statt ökonomischer Effizienzkriterien fordern Akteur:innen eine diskursive Ausgestaltung schulischer Entwicklungsprozesse, um gesellschaftliche Teilhabe und demokratische Mitbestimmung zu gewährleisten (Die Linke, BT_195b, 2019, S. 10). Die GEW unterstützt diese Position und hebt hervor, Digitalisierung dürfe nicht „zum Einfallstor der Profitinteressen von Unternehmen“ (GEW_1275, 2021, S. 2) werden.

Dabei wird im Zuge der staatsbürgerlichen Kritik auch die mit der industriellen Logik verbundene Absicht der Leistungs- und Verhaltensmetrisierung hinterfragt. Die GEW warnt, dass „nicht erst die Auswertung [sic!] sondern bereits die systematische und standardisierte Erfassung von Leistungs- und Verhaltensstrukturen von Anwender*innen – insbesondere Kindern und Jugendlichen – als problematisch betrachtet“ (GEW_1338, 2021, S. 24) werden müsse. Diese Metrisierung greife tiefer als bloße Bewertungspraxis, sie verändere den Bildungsbegriff selbst. „Lernerfolg ist eben nicht ein Ergebnis automatisierter Prozesse, und das Ziel von Lernen ist nicht Messbarkeit“, so Stumpp (Die Grünen, BT_244c, 2019, S. 14551). Darüber hinaus kritisiert die GEW, dass Learning Analytics nicht der Förderung von Selbstständigkeit, sondern der „Perfektionierung des taxierenden Systems der Ziffernnoten und Vergleichsarbeiten“ (GEW_1265, 2020, S. 2) diene, während das HFD ergänzt, die Komplexität von Bildungsprozessen lasse sich nicht in algorithmischen Mustern erfassen, weshalb es einer bildungstheoretischen Reflexion bedürfe, die nicht von informatischen Disziplinen dominiert werde (HFD_1202, 2021, S. 52). Der viel zitierte britische Abiturnotenskandal habe zudem gezeigt, welche gesellschaftlichen Spannungen algorithmische Bewertungssysteme erzeugen können. Schülerinnen und Schüler protestierten gegen eine maschinelle Benotung, weil sie „als

Individuen und nicht als statistisches Zahlenmaterial“ (GEW_1268, 2020, S. 1) wahrgenommen werden wollten.

Vor diesem Hintergrund stellen verschiedene Akteur:innen die Legitimität der industriellen Effizienzkriterien als Bildungsziele infrage. Zum einen sei zu klären, ob Effizienz und Leistungsfähigkeit überhaupt legitime Bildungsziele sein können (HFD_1191, 2016, S. 18, 27 f.). Zum anderen drohen zentrale Herausforderungen wie die Heterogenität der Lernenden oder das Ziel inklusiver Bildung aus dem Blick zu geraten, wenn der Diskurs allein auf Optimierung und Output fokussiert bleibt (Stumpp/Die Grünen, BT_244c, 2019, S. 14551). In dieser Perspektive erscheint die industrielle Rechtfertigungsordnung als Bedrohung demokratischer und pädagogischer Werte, weil sie Lernende zu Datensubjekten und Lehrende zu Effizienzverwalter:innen reduziert.

Die dargestellten Spannungen zwischen humanistischem Bildungsverständnis und den industriellen wie solutionistischen Ziellogiken bleiben dabei jedoch nicht auf der Ebene normativer Werteaushandlung verhaftet. Sie materialisieren sich in institutionellen Strukturen, in denen wirtschaftliche Akteur:innen zunehmend Einfluss auf Bildungspolitik und pädagogische Praxis ausüben. Damit verschiebt sich die Existenzprüfung des humanistischen Bildungsideals von einer moralischen auf eine organisatorische Ebene. Staatsbürgerliche Kritik richtet sich folglich gegen die technokratische Steuerung durch solutionistische Akteur:innen ebenso wie gegen die industriepolitische Einflussnahme auf Bildungsinstitutionen. Diese beiden Einflussformen werden im Folgenden gemeinsam betrachtet, da sie Ausdruck einer strukturellen Verschiebung sind, in der ökonomische Rationalitäten die Gestalt von Bildung prägen.

Die staatsbürgerliche Kritik des Einflusses wirtschaftlicher Interessen und technokratischer Steuerung auf das Bildungssystem

Die zuvor dargestellten Spannungen zwischen humanistischem Bildungsideal und den industriellen wie solutionistischen Ziellogiken verweisen auf eine strukturelle Dimension: den zunehmenden Einfluss wirtschaftlicher Akteur:innen auf die politische und pädagogische Gestaltung von Bildung. Die moralischen Konflikte zwischen Effizienz, Innovation und Emanzipation übersetzen sich dabei in institutionelle Machtverhältnisse, die sowohl die Steuerung von Bildungsprozessen als auch die Definition dessen betreffen, was als „gute Bildung“ gilt.

In der bildungspolitischen Debatte wird der Einsatz von KI, insbesondere von Learning Analytics, mitunter als technokratischer Eingriff verstanden, der kritische und individuelle Lernprozesse zugunsten standardisierter Datenanalysen verdrängt (Stumpp/Grüne, BT_244c, 2019, S. 14551). Die GEW kritisiert, dass staatliche und kommerzielle Interessen zunehmend in

einem undurchsichtigen Netz verflochten seien, um technokratische Lösungen durchzusetzen. Diese Verflechtungen führten zu wachsender Abhängigkeit von wirtschaftlichen Interessen, während gleichzeitig demokratische Partizipationsversprechen als legitimierender Deckmantel neoliberaler Bestrebungen der Kommerzialisierung des Bildungswesens dienten (GEW_1244, 2019, S. 16). So betrifft ein zentraler Kritikpunkt die zunehmende Privatisierung bildungstechnologischer Infrastrukturen. Dabei problematisiert das HFD, „dass die Programmierung der Algorithmen in privater Hand liegt und sich einer demokratischen Kontrolle und Legitimation entzieht“ (HFD_1191, 2016, S. 37). Diese Abhängigkeit verschiebt Verantwortung von öffentlichen Institutionen zu privatwirtschaftlichen Akteur:innen. Personalisierte Lernangebote seien daher Ausdruck „einer bestimmten Konstellation ökonomischer und politischer Interessen“ (ebd.), in der pädagogische Ziele ökonomischen Logiken untergeordnet werden. Ein Interessenkonflikt, der auch von der Abgeordneten Bull-Bischoff (Die Linke) aufgegriffen wird:

> Digitale Lernumgebungen müssen das Werkzeug der Lernenden und Lehrenden werden oder bleiben, statt beides aus der Cloud heraus algorithmisch berechnet zu steuern. Es kann gut sein, dass dieses Ziel mit den Interessen großer IT-Konzerne und -Lobbyisten und den Absichten, die sie mit Learning Analytics verbinden, nicht zusammengeht. Deshalb bräuchten Pädagogen sowie Lernende Aufmerksamkeit, kritische Distanz und die Besinnung darauf, was Bildung eigentlich bedeutet und was Bildung eigentlich braucht. (BT_244c, 2019, S. 14627 f.)

Diese Kritik richtet sich sowohl gegen solutionistische Ansprüche, Bildung durch datengetriebene Optimierung zu verbessern, als auch gegen industrielle Effizienzlogiken, die Lernprozesse nach Produktivitätsprinzipien bewerten. Aus staatsbürgerlicher Sicht droht damit zum einen eine Entpolitisierung von Bildung: Wo privatwirtschaftlich gestaltete Algorithmen und Kennzahlen die Steuerung übernehmen, wird demokratische Aushandlung durch technische und ökonomische Verfahren ersetzt. Zum anderen wirft dies Fragen zur sozialen Konstruktion von Technik und zu inkorporierten Machtverhältnissen auf (HFD_1191, 2016, S. 18). Sitte (Die Linke) warnt vor einer „ungebremste[n] Wachstumsideologie“ (BT_113d, 2020, S. 23760), die soziale Macht- und Wohlstandsungleichheiten verstärke. Daran anknüpfend weist die GEW bezugnehmend auf die Soziologin Hartong darauf hin, dass KI-Systeme nie neutral seien, sondern die Werte und Interessen ihrer Entwickler:innen widerspiegeln (GEW_1239, 2019, S. 1).

Das HFD weist zudem darauf hin, dass Unternehmen und Stiftungen im Spannungsfeld zwischen dem Anspruch, mündige Bürger:innen auszubilden, und dem Ziel, Fachkräfte für den Arbeitsmarkt bereitzustellen, gezielt durch finanzielle Anreize Einfluss auf Forschung und

Lehre nehmen (HFD_1202, 2021, S. 46). Es erkennt darüber hinaus eine „Vereinnahmung beziehungsweise Instrumentalisierung eines pädagogischen Konzepts durch ökonomische Interessen der Technologiehersteller (…). Denn die als philanthropisch ausgegebenen Ziele offenbaren bei genauerem Hinsehen eine den klassischen pädagogischen Idealen entgegenstehende Wirkungsweise“ (HFD_1191, 2016, S. 37). Diese Form der sogenannten Individualisierung führe bei genauerem Hinsehen zur Entmündigung der Lernenden und zur Verengung der Lerninhalte auf prüfungsrelevante Aspekte (ebd.). Damit kritisiert das HFD explizit den Versuch, solutionistische Interessen unter dem Deckmantel staatsbürgerlicher Ideale zu verfolgen. Vor diesem Hintergrund verweist GEW-Vorstandsmitglied Klinger auf die Notwendigkeit störfreier Räume, in denen „unbeobachtetes Lernen und Arbeiten möglich ist“ (GEW_1239, 2019, S. 2). Weiterhin fordert die Gewerkschaft eine klare Begrenzung wirtschaftlicher Einflussnahme:

> Hard- und Softwarekonzerne dürfen keinen Einfluss auf Unterrichtsinhalte nehmen. (…) Auch GEW-Schulexpertin Ilka Hoffmann betont, digitale Materialien im Netz dürften ‚nicht zu einem weiteren Einfallstor für die Kommerzialisierung schulischer Bildung werden‘. Der Einfluss von Privatwirtschaft und Lobbyisten müsse begrenzt werden. ‚Bildung ist ein öffentliches Gut und kein Geschäftsmodell.‘ (GEW_1285, o. J., S. 2)

Insgesamt zeigt sich, dass die wirtschaftliche Einflussnahme auf Bildung eine doppelte Existenzprüfung des humanistischen Bildungsideals hervorruft. Solutionistische Akteur:innen gefährden dieses Ideal durch technokratische Steuerung und algorithmische Intransparenz, während industrielle Akteur:innen es durch Wachstums- und Effizienzlogiken unterminieren. Gemeinsam verschieben sie den normativen Maßstab von Bildung weg von Emanzipation und Teilhabe, hin zu Steuerbarkeit und Verwertbarkeit. Die staatsbürgerliche Kritik, wie sie in den untersuchten Debatten formuliert wird, kann somit als Versuch gelesen werden, die Grenzen technischer und ökonomischer Rationalität im Bildungswesen neu auszuhandeln. Bildung erscheint als Prüfstein demokratischer Selbstbestimmung, deren Schutz vor Kommerzialisierung und technokratischer Vereinnahmung zur Voraussetzung legitimer KI-Governance wird.

9.5.1.3. Die Verletzung von Persönlichkeitsrechten durch den Einsatz Künstlicher Intelligenz

Die Frage nach dem Schutz von Persönlichkeitsrechten und der Privatheit bildet einen weiteren zentralen Bezugspunkt staatsbürgerlicher Kritik am KI-Einsatz im Bildungsbereich und stellt damit zugleich Ansprüche der Cité der Solution, der Industrie und des Marktes infrage. In den untersuchten Dokumenten richtet sich diese Kritik wiederkehrend gegen umfassende Datensammlungen, Profilbildungen und intransparente algorithmische Entscheidungen, die als Eingriffe in Selbstbestimmung und Privatsphäre verstanden werden (s. z. B. WR_1177, 2022,

S. 31 ff.). Im Kontext von Learning Analytics wird etwa gewarnt, „dass das Datenprofil hinter einer Person wichtiger wird als die Person selbst“ (Völlers/SPD, BT_244c, 2019, S. 14626). Dies führe schließlich zu einem „gläsernen Schüler“, so die existenzielle Kritik Espendillers von der AfD-Fraktion, weshalb Daten gar nicht erst erhoben werden sollten (BT_244c, 2019, S. 14548 f.). Diese Kritik wird im Kontext der KI-basierten Weiterbildungsplattform „Milla“ durch die GEW auf den „gläsernen Arbeitnehmer“ (GEW_1250, 2019, S. 1) ausgeweitet. Kritik entfaltet sich zudem an mangelnder Transparenz und Verantwortungszuordnung bei Datenerhebung, -nutzung und algorithmischer Verarbeitung, wodurch die Souveränität des Einzelnen eingeschränkt wird (Sitte/Die Linke, BT_113d, 2020, S. 23760; BT_244c, 2019, S. 14627; GEW_1244, 2019, S. 13; WR_1143, 2020, S. 12). Ebenso werden unklare Eigentums- und Verantwortlichkeitsfragen der Erhebung von Lernendendaten kritisiert (HFD_1188, 2016, S. 2).

Diese staatsbürgerliche Kritik trifft nun auf drei unterschiedliche Rechtfertigungslogiken, die jeweils eine eigene Existenzprüfung auslösen: die solutionistische Hoffnung auf technische Problemlösung, die industrielle Orientierung an Effizienz und die marktorientierte Wettbewerbslogik. Darüber hinaus verschärft sich die Debatte im Kontext eines solutionistischen KI-Einsatzes, indem die staatsbürgerliche Kritik der eingeschränkten Persönlichkeitsrechte auf grundlegende Überwachungsbefürchtungen durch den KI-Einsatz ausgeweitet wird.

Der Konflikt zwischen solutionistischer Problemlösung und staatsbürgerlichen Datenschutzansprüchen

Zunächst richtet sich die staatsbürgerliche Kritik im Konflikt mit der Cité der Solution jedoch gegen die Tendenz, Datenschutzrisiken als nachrangig gegenüber vermeintlichen Lerngewinnen zu behandeln. Die Debatte spitzt sich an einer Gratwanderung zu, einerseits

> (…) den Schülerinnen und Schülern die bestmögliche Bildung zur Verfügung zu stellen, was die Ausschöpfung der Potenziale von Learning Analytics einbezieht. Zugleich haben Bund und Länder die Pflicht, die individuellen Lerndaten von Schülerinnen und Schülern vor unbefugtem Zugriff zu schützen. Denn von beidem hängen in einer digitalen und globalisierten Welt künftige Bildungs-, Karriere- und Lebenschancen insgesamt ab. (BT_244a, 2019, S. 2)

Kontroversen entzünden sich an intelligenten Tutorensystemen wie der Anwendung Area9 Rhapsode und an KI-basierten Prüfungsformaten, denen eine Bedrohung des Datenschutzes attestiert wird (LT-SN_1014, 2022, S. 2 f.; WR_1177, 2022, S. 31). Parteienübergreifend wird insbesondere die von der FDP vertretene solutionistische Forderung eines flächendeckenden Learning-Analytics-Einsatzes kritisiert, welcher adäquate Datenschutzstandards ausblende

(BT_244b, 2020, S. 8; Espendiller/AfD, BT244c, 2019, S. 14548 f.; Röspel/SPD, BT244c, 2019, S. 14559).

Zwar verweisen Befürworter:innen auf einen hohen Umsetzungswillen der Lehrkräfte, Lernendendaten zu nutzen, und fordern Rechtssicherheit bezüglich der Datenschutzregularien (Suding/FDP, BT_244c, 2019, S. 14547). Aus staatsbürgerlicher Sicht bleibt jedoch problematisch, dass die Datenschutzdebatte, ebenso wie ethische Leitlinien, häufig abstrakt bleibt und die Grundfrage nach Überwachung, Profiling und Persönlichkeitsrechten nicht hinreichend adressiert (GEW_1244, 2019, S. 18). So finden sich in der bildungspolitischen Debatte etwa diffuse Forderungen nach nicht weiter spezifizierten Maßnahmen, um „bei der Sammlung von Daten die Datensouveränität des Bürgers zu wahren“ (BT_157b, 2019, S. 6). Dieses Spannungsfeld zwischen staatsbürgerlicher Kritik und solutionistisch gerechtfertigter Datennutzung wird exemplarisch durch das TAB skizziert:

> Learning Analytics bewegen sich im Spannungsfeld zwischen rechtlichen und ethischen Bedenken einerseits und (berechtigtem) Erkenntnisinteresse von Lehrenden, Lernenden, Bildungseinrichtungen und Bildungsforschenden sowie den technischen Möglichkeiten der umfangreichen Datenerhebung und -auswertung andererseits. Generell wird eine Diskrepanz zwischen dem technisch Möglichen und den bestehenden Regulierungen deutlich. (TAB_1335, 2020, S. 223).

Vor dem Hintergrund dieses Spannungsverhältnisses sei daher „eine grundlegende Verständigung über die Ziele und Reichweite solcher [KI-]Verfahren, die jenseits von Technikphobie und Bildungsrevolution angesiedelt sein sollte“ (HFD_1192, 2016, S. 3), erforderlich, so das HFD. Diese nivellierende Perspektive hebt somit den Konflikt zwischen solutionistischer Problemlösung und staatsbürgerlichen Ansprüchen der Wahrung von Persönlichkeitsrechten hervor. Aus staatsbürgerlicher Perspektive steht nicht die Nützlichkeit einzelner Tools im Vordergrund, sondern die Grenzziehung durch Grundrechte. Datenschutz erscheint hier als Bedingung der Möglichkeit legitimer Innovation, nicht als nachträgliche Compliance-Aufgabe.

Die staatsbürgerliche Kritik an algorithmischer Überwachung und solutionistischen Kontrolllogiken

Die staatsbürgerliche Kritik an der Verletzung von Persönlichkeitsrechten erfährt in der Debatte um Überwachung eine deutliche Zuspitzung. Während der Konflikt bis hierhin den Schutz individueller Daten und der Privatsphäre betont, richtet sich der Fokus hier auf die Gefahr einer tiefgreifenden Veränderung der sozialen Ordnung durch die systematische Erfassung, Auswertung und Steuerung menschlichen Verhaltens mittels KI. Diese Entwicklung wird als fundamentale Bedrohung demokratischer Grundrechte wahrgenommen (BT_113b, 2020, S. 324;

Sitte/Die Linke, BT_113d, 2020, S. 23760; BT_157a, 2018, S. 2; Stumpp/Die Grünen, BT_244c, 2019, S. 12551; GEW_1269, 2020, S. 1; GEW_1338, 2021, S. 28). Die staatsbürgerliche Kritik der Überwachung verknüpft sich mit der Sorge vor einer zunehmenden Quantifizierung des Sozialen. Die umfassende Datenerhebung, etwa im Rahmen von Learning Analytics, wird nicht nur als Eingriff in die Privatsphäre, sondern mit einer Stärkung des Überwachungsstaats in Verbindung gebracht (GEW_1250, 2019, S. 1). In den Dokumenten finden sich wiederholt Warnungen vor „dystopische[n] Wirkungsszenarien" (GEW_1244, 2019, S. 17) und „totalitäre[n] Tendenzen durch Big Data Analysen und Manipulationen durch eine technokratische Verhaltens- und Gesellschaftssteuerung" (GEW_1338, 2021, S. 28). Hier zeigt sich, dass die Überwachungskritik eine radikale Form annimmt: Sie zielt nicht mehr nur auf einzelne Anwendungen, sondern auf die normative Ordnung einer demokratischen Gesellschaft und stellt den KI-Einsatz grundsätzlich infrage.

In diesem Kontext geraten solutionistische Rechtfertigungen in einen grundlegenden Wertekonflikt mit staatsbürgerlichen Idealen. Anwendungsbeispiele wie die Prognose von Studienabbrüchen oder die in chinesischen Klassenzimmern vorgenommene Überwachung von Schüler:innen lösen starke Bedenken aus und setzen solutionistischen Rechtfertigungen Grenzen, indem entsprechende Einsatzszenarien abgelehnt werden, da sie staatsbürgerlichen Idealen entgegenstehen (BT_113b, 2020, S. 334). Gerade das Szenario einer KI-basierten Überwachung in chinesischen Klassenzimmern lässt die AfD „Orwell'sche" Zustände heraufbeschwören (Jongen/AfD, BT_113d, 2020, S. 23763; BT_244b, 2020, S. 7) und auch die GEW wirft in Anbetracht flächendeckender Leistungserfassung durch Learning Analytics die plakative Frage auf: „Ist 2021 das neue 1984?" (GEW_1338, 2021, S. 24).

Dabei wird die Priorisierung solutionistischer Steuerungs- und Problemlösungslogiken kritisiert, die in Verbindung mit einer intransparenten Datensammlung und automatisierten Kontrolle zu einer Erosion individueller Persönlichkeitsrechte führen und dadurch mit dem Ideal demokratischer Selbstbestimmung unvereinbar seien (GEW_1244, 2019, S. 19). Damit wird der Konflikt zwischen staatsbürgerlicher und solutionistischer Logik als eine moralische Grenzziehung sichtbar: Während die Cité der Solution den KI-Einsatz als instrumentell beherrschbares Mittel der Problemlösung im Bildungssektor interpretiert, erkennt die Cité des Staatsbürgers darin eine Bedrohung der Freiheit und die Gefahr gesellschaftlicher Fremdsteuerung.

Während in dieser Kritik eine fatalistisch-deterministische Haltung zum Ausdruck kommt, fordert Bayaz (Die Grünen), den technischen Wandel politisch und normativ zu gestalten, statt ihn

faktisch zu akzeptieren, und formuliert damit eine Gegenposition mit Blick auf die Gefahr technikdeterministischer Pfadabhängigkeiten:

> Wenn eine neue Technologie aufkommt, dann wirkt ja manchmal die normative Kraft des Faktischen. Es ist eine Wette – diese Wette kann gut ausgehen; sie kann aber auch schlecht ausgehen –, und ich finde, wir sollten uns dieser Wette nicht einfach ausliefern. Nur wenn wir selbst gestalten, sind wir in der Lage, Normen zu setzen: Normen, an denen sich andere orientieren, vielleicht sogar orientieren müssen; Normen, die mit unseren Werten vereinbar sind. (Bayaz/Die Grünen, BT_113d, 2020, S. 23765)

Die Überwachungskritik fungiert somit als radikale Kritik des KI-Einsatzes und als verschärfte Existenzprüfung: Sie markiert den Punkt, an dem solutionistische KI-Ansprüche nicht nur auf Fragen der Legalität, sondern auf den Kern demokratischer Selbstbestimmung treffen. Die Cité des Staatsbürgers behauptet ihre moralische Größe durch das Beharren auf Freiheit, Transparenz und Mitbestimmung als unantastbare Grenzen technischer Steuerung.

Der Konflikt zwischen staatsbürgerlichem Grundrechtsschutz und industriellem Innovationsstreben

Der Konflikt zwischen einem staatsbürgerlichen Schutz der Grundrechte und der industriellen Effizienzorientierung äußert sich hingegen weniger dystopisch. Stattdessen wird diese Spannung vielmehr als zentrale Gestaltungsaufgabe der Bildungspolitik gerahmt. Politische Akteur:innen betonen, dass es darum gehe, „die Entwicklung und Nutzung innovativer KI-Systeme bei gleichzeitiger Wahrung der Sicherheit und der Grundrechte zu fördern“ (LT-SH_934, 2021, S. 11). Diese Ambivalenz wird als erhebliche politische Herausforderung anerkannt: KI-basierte Lehr- und Lernmittel versprechen Effizienz, zugleich drohen jedoch Einschränkungen von Privatsphäre, Selbstbestimmung und Chancengleichheit (BT_113b, 2020, S. 297; Durz/Union, BT_113c, 2018, S. 4162; Völlers/SPD, BT_244c, 2019, S. 14626; HFD_1191, 2016, S. 27). Aus staatsbürgerlicher Sicht entsteht ein struktureller Zielkonflikt, wenn Effizienzgewinne über datenintensive Praktiken realisiert werden, deren Voraussetzungen gerade jene Persönlichkeitsrechte etwa von Schüler:innen tangieren, die als besonders schützenswert gelten. Entsprechend kritisiert die Grünen-Fraktion einen „unkritischen Fortschrittsglauben“ (BT_244b, 2020, S. 8), der Datenschutzrisiken verharmlost, und betont die Bedeutung der Schulen als „besonderer Schutzraum“ (ebd.). Zuspitzungen wie die Reduktion von Lernenden und Lehrkräften auf „reine Datenlieferanten“ (Espendiller/AfD, BT_244c, 2019, S. 14549) verdeutlichen das Wahrnehmungsmuster einer instrumentellen Vereinnahmung personenbezogener Daten zugunsten industrieller Optimierungsziele. Im Unterschied zum solutionistischen Konflikt, der den pädagogischen Nutzen problematisiert, richtet sich die staatsbürgerliche

Kritik hier gegen eine produktionslogische Rahmung von Bildung als Effizienzsystem. Die Existenzprüfung betrifft die Frage, ob Effizienz und Produktivität als legitime Bildungsziele gelten können, wenn ihre Umsetzung strukturell auf eine Erosion individueller Rechte angewiesen ist (Die Linke, BT_244b, 2020, S. 8; HFD_1206, 2021, S. 2; WR_1177, 2022, S. 39).

Zugleich legt die Analyse wechselseitige Kritikmuster offen. So werden staatsbürgerliche Leitprinzipien wie hohe Datenschutzstandards aus industrieller Sicht als strukturelle Hürden markiert. Sie erschwerten etwa die Einführung von Learning Analytics oder KI-gestützten Prüfungsformaten, da Transaktionskosten, Urheberrechte und datenschutzrechtliche Auflagen die Umsetzung erheblich verzögerten (AfD, BT_157b, 2019, S. 6; BT_244a, 2019, S. 1 f.; TAB_1335, 2020, S. 180; 242). Auch die EFI betont, dass der Einsatz von Lern- und Analysesoftware Potenzial zur Verbesserung von Lehrqualität und Ressourcennutzung biete, jedoch zunächst Datenschutzfragen zu klären seien (EFI_1082, 2019, S. 103).

In dieser wechselseitigen Kritik wird die Regulierung zur zentralen Streitfrage. Während die Cité des Staatsbürgers rechtliche Sicherheit als Voraussetzung legitimer Technikgestaltung fordert, betonen Akteur:innen, die aus der Sicht der Cité der Industrie argumentieren, die umgekehrte Kausalität: Erst klare Erlaubnistatbestände und ein innovationsfreundlicher Rechtsrahmen machten einen verantwortungsvollen Einsatz überhaupt möglich. Die FDP kritisiert in diesem Zusammenhang, dass „häufig nur klar [ist], was verboten ist, nicht aber, was tatsächlich möglich ist“ (BT_244a, 2019, S. 2). Auch die Enquete-Kommission erkennt in ihrer Abwägung zwischen Chancen und Risiken, dass staatsbürgerliche Beteiligungsrechte, wie etwa Mitbestimmung bei der Einführung von KI, Innovationsfortschritt verlangsamen können, wenngleich diese auch konstruktiv wirken können (BT_113b, 2020, S. 350).

Trotz der divergierenden Bewertungen besteht weitgehende Einigkeit darüber, dass ein verlässlicher Rechtsrahmen für den Einsatz von KI im Bildungsbereich notwendig ist, der sowohl Datenschutz als auch Innovation berücksichtigt (BT_113b, 2020, S. 34; WR_1177, 2022, S. 80). Der Unterschied liegt im normativen Schwerpunkt: Aus staatsbürgerlicher Sicht hat der Schutz individueller Rechte Vorrang, während die Industrie Legitimität über die Ermöglichung von Innovation und Effizienz definiert. Diese Divergenz scheint jedoch vorerst ungelöst zu bleiben. Aus Sicht der FDP bedarf es sowohl gezielter Investitionen in Hochschulbildung und Forschung als auch „politische Rahmenbedingungen, die die Nutzung von Daten in der Schule nicht verhindern, sondern verantwortungsvoll ermöglichen, um die Qualität des Bildungswesens insgesamt zu verbessern“ (BT_244b, 2020, S. 1). Auch die US-amerikanische Neurowissenschaftlerin Feldmann Barrett spricht sich in einem TAB-Bericht gegen ein generelles Verbot

umstrittener KI-Anwendungen wie der Emotionserkennung aus und plädiert für eine Stärkung wissenschaftlicher Forschung, um bestehende Unsicherheiten zu verringern und Chancen zu nutzen (TAB_1327, 2021, S. 5). Ebenso fordert die niederländische Hochschulinitiative SURFnet, ethische und rechtliche Grenzen stärker zu berücksichtigen, ohne das Potenzial von Learning Analytics grundsätzlich infrage zu stellen (HFD_1190, 2016, S. 4). Positionen wie diese erkennen die Unsicherheit der Technikgestaltung und -wirkungen an, ohne daraus jedoch ein Anwendungshemmnis abzuleiten (Bartol/SPD, BT_113c, 2018, S. 4150).

Insgesamt zeigt sich eine wechselseitige Existenzprüfung: Für die Cité des Staatsbürgers steht Legitimität unter dem Vorzeichen des Schutzes von Grundrechten, für die Cité der Industrie unter dem der Ermöglichung von Innovation. Beide Perspektiven verweisen mit der Kritik einer fehlenden Klarheit über rechtliche Rahmenbedingungen auf denselben Engpass, auf den nachfolgend noch einmal eingegangen wird (siehe Abschnitt 9.5.1.4).

Der Konflikt zwischen staatsbürgerlicher Gemeinwohlorientierung und marktgetriebener Wettbewerbslogik

Der Konflikt mit der Cité des Marktes verschiebt den Fokus auf Wettbewerbs- und Kommerzialisierungslogiken. Staatsbürgerliche Kritik richtet sich gegen die Priorisierung von Standort- und Wettbewerbsvorteilen gegenüber Gemeinwohlorientierung und Grundrechtsschutz (Sitte/Die Linke, BT_113d, 2020, S. 23760). Zugleich betont Schüle (SPD), „der Mensch und nicht der Markt“ (BT_157c, 2018, S. 8100) müsse im Zentrum der KI-Entwicklung stehen; internationaler Wettbewerbsdruck dürfe nicht zu Lasten des Datenschutzes gehen (Schüle/SPD, BT_157d, 2019, S. 9473). Bemerkenswert ist die Ambivalenz in dieser Argumentation: So sei nur zu berücksichtigen, „dass die Menschen selbstbestimmt an der digitalen Welt von heute und morgen teilhaben können, sondern auch (…), dass die vielen Vorteile der neuen Schlüsseltechnologien ihren Eingang in den Markt finden“ (Schüle/SPD, BT_157c, 2018, S. 8100). Damit knüpft sie staatsbürgerliche Forderungen an Marktlogiken an, versucht sie jedoch durch demokratische Werte zu begrenzen und auszubalancieren.

Konkrete Kritikpunkte betreffen die Kommodifizierung von Bildungsdaten. Selbst aus marktliberaler Perspektive der FDP wird die Notwendigkeit betont, „eine Vermarktung von Kita-, Schul- oder Hochschulprofilen von Schülerinnen und Schülern, aber auch deren Lehrkräften und Eltern, zu verhindern“ (BT_244a, 2019, S. 4). Gleichzeitig erzeugt das durch die Cité des Marktes gerechtfertigte Individualisierungsversprechen durch KI-gestützte, personalisierte Lernangebote neue Transparenz- und Datenschutzbedenken (LT-NW_753, 2020, S. 179). Offene Fragen zu Dateneigentum, institutionellen Beziehungen und Übergängen zwischen

Bildungssektoren verstärken die staatsbürgerliche Skepsis gegenüber marktgetriebenen Datenflüssen und möglichen Zweitverwertungen (HFD_1188, 2016, S. 2).

Die Analyse belegt jedoch, dass sich dieser Konflikt in einer wechselseitigen Kritik verdichtet. Aus marktorientierter Perspektive wird wiederholt auf erfolgreiche KI-Konzepte im Ausland verwiesen, die als Referenz für Wettbewerbsfähigkeit gelten. Die FDP betont, in Ländern wie den USA, China, Japan oder Estland sei KI im Bildungsbereich bereits Alltag; Deutschland riskiere, den Anschluss zu verlieren (BT_244b, 2020, S. 1; Suding/FDP, BT_244c, 2019, S. 14547). Dabei stünden jedoch Datenschutzbedenken und unzureichende oder mitunter auch als zu strikt bewertete rechtliche Rahmenbedingungen beispielsweise einem erfolgreichen Einsatz von Learning Analytics, der als Hebel für individualisiertes Lernen und Qualitätssteigerung betrachtet wird, im Wege (FDP, BT_244b, 2020, S. 4; Knoerig/Union, BT_251f, 2020, S. 18406; TAB_133, 2020, S. 144). Die Zuspitzung des AfD-Abgeordneten Komning formuliert die Sorge, KI-Verbote würden „Deutschland an das Ende der Liste der wettbewerbsfähigen Länder [katapultieren]“ (BT_251f, 2020, S. 18406).

Die marktorientierte Existenzprüfung richtet sich auf die internationale Vergleichsordnung. Staatsbürgerliche Leitplanken erscheinen hier weniger als Schutz der Freiheit, sondern als Hemmnis wirtschaftlicher Dynamik. Der Bewertungsmaßstab verschiebt sich von Grundrechten zu Standort- und Zeitvorteilen. Damit gilt die staatsbürgerliche Forderung nach Regulierung und Gemeinwohlorientierung als potenzielles Wettbewerbshindernis, während umgekehrt der marktorientierte Anspruch durch den Ruf nach demokratischer Kontrolle und Persönlichkeitsrechten begrenzt wird. Der Konflikt unterscheidet sich von den vorherigen durch die zentrale Rolle von Kommerzialisierung und Wettbewerb. Im Fokus stehen die Legitimität der Verwertung personenbezogener Bildungsdaten und die Gefahr, dass Marktmechanismen Grundrechte untergraben. Insgesamt zeigen sich drei Konfliktlinien: zwischen solutionistischem KI-Einsatz und Datenschutzansprüchen (Solution), zwischen Effizienzstreben und Grundrechtsschutz (Industrie) sowie zwischen Wettbewerbsorientierung und Gemeinwohl (Markt). Gemeinsam ist allen drei Fällen, dass Transparenzappelle und allgemeine Ethikleitlinien die Grundrechtsproblematik bislang nicht hinreichend auflösen. Die empirischen Dokumente deuten vielmehr darauf hin, dass Legitimation an vorgelagerte Klärungen zu Zielen, Datengrenzen, Verantwortlichkeiten und Aufsicht gebunden ist, bevor technische Potenziale ausgeschöpft werden können (TAB_1335, 2020, S. 223; HFD_1192, 2016, S. 3; GEW_1244, 2019, S. 18).

9.5.1.4. Unsicherheit über Technikfolgen und Regulierungsbedarfe als Prüfstein staatsbürgerlicher Legitimation

Die zuvor rekonstruierten Konflikte zwischen der Cité des Staatsbürgers und den Cités der Solution, der Industrie und des Marktes verweisen auf eine gemeinsame Grundspannung der Unsicherheit über die rechtlichen, normativen und epistemischen Bedingungen des KI-Einsatzes im Bildungswesen. Für die Cité des Staatsbürgers wird diese Unsicherheit zu einer Realitätsprüfung, da Legalität und Regelklarheit zentrale Bestandteile ihrer moralischen Größe darstellen.[34]

Wie die vorhergehenden Analysen zeigen, besteht eine doppelte Form von Unsicherheit. Einerseits bleibt umstritten, welche Werte und Normen die Legitimation des KI-Einsatzes bestimmen sollen, beispielsweise im Spannungsfeld zwischen Chancengleichheit, Datenschutz und Innovationsförderung. Andererseits herrscht wissenschaftliche Unklarheit über die langfristigen, teils irreversiblen Folgen des Technologieeinsatzes (TAB_1333, 2021, S. 5). Diese Kombination aus normativer und epistemischer Unsicherheit mündet zum einen in der staatsbürgerlichen Kritik des voreiligen solutionistischen Technikeinsatzes. Zum anderen lässt sie die Frage der Technikregulierung selbst zu einem Prüfstein demokratischer Legitimation und Steuerungsfähigkeit werden. So betont etwa Katja Suding (FDP): „Wir brauchen endlich Klarheit darüber, was erlaubt ist und was nicht“ (BT_244c, 2019, S. 14547). Die Debatte um Technikregulierung wird so zur Realitätsprüfung des staatsbürgerlichen Legalitätsanspruchs. Sie ist zugleich als Reaktion auf Kritik und als Vorsorge gegen unkontrollierte technische Dynamiken zu deuten.

Die staatsbürgerliche Kritik mangelnder Evidenz als Risiko eines solutionistischen Einsatzes Künstlicher Intelligenz und als Herausforderung der Technikregulierung

Besonders deutlich tritt diese Unsicherheit im Konflikt zwischen der staatsbürgerlichen und der solutionistischen Cité zutage. Aus staatsbürgerlicher Sicht wird der flächendeckende, solutionistisch gerechtfertigte Einsatz von KI kritisiert, da er oftmals auf unzureichend geprüften Annahmen über Nutzen und Wirksamkeit beruht. Brünler (Die Linke) bemängelt, dass der Ausbau von KI-Anwendungen unabhängig von empirischen Erkenntnissen vorangetrieben werde. Geplante Evaluationsmaßnahmen fänden zu wenig Beachtung, während widersprüchliche Ergebnisse teilweise ausgeblendet würden (LT-SN_904, 2021, S. 2788). Somit wird die bestehende

[34] An dieser Stelle wird die Rolle der Unsicherheit am Beispiel der Technikregulierung verdeutlicht, die den Legalitätsansprüchen der Cité des Staatsbürgers zugeordnet werden kann. Gleichzeitig fungiert die (wissenschaftliche) Unsicherheit auch in anderen Cités als Realitätsprüfung. So wurden Forderungen nach Rechtssicherheit bereits im Rahmen verschiedener Rechtfertigungslogiken, insbesondere in staatsbürgerlichen Kompromissen, aufgegriffen (siehe Abschnitt 9.5.1.3). Darüber hinaus wird Unsicherheit beispielsweise für die Cité des Hauses zu einer Realitätsprüfung aufgrund der unbekannten Folgen des KI-Einsatzes, etwa in Bezug auf pädagogische Konsequenzen (z. B. BT_244b, 2020, S. 8).

Unsicherheit nicht verringert, sondern verdeckt. Christmann (Die Grünen) fordert daher aufgrund der potenziellen Gefahren einerseits und der zahlreichen solutionistisch-wünschenswerten Einsatzbereiche andererseits eine sorgfältige Chancen-Risiko-Abwägung und kritische Bewertung der zu erreichenden Ziele (BT_157c, 2018, S. 8096). Wenngleich auch die Bundesregierung betont, dass Chancen und Risiken des KI-Einsatzes im Bildungswesen sorgsam gegeneinander abzuwägen seien (BT_244b, 2020, S. 9), kritisiert Christmann die bisherigen Bestrebungen als unzureichend:

> Eine Technologie, die so viel kann, wirft natürlich auch viele Fragen auf und bringt auch potenzielle Gefahren mit sich. (…) Deswegen ist es umso wichtiger, dass wir sie, um sie zu unserem Wohl zu nutzen, exzellent beherrschen und dass wir eine klare Vision haben, für welche Zwecke wir sie einsetzen wollen, und dafür braucht es deutlich mehr Anstrengungen dieser Bundesregierung als bisher. (BT_157c, 2018, S. 8096)

Der aus staatsbürgerlicher Perspektive zentrale Kritikpunkt besteht darin, dass solutionistische Erwartungen politisch verfestigt werden, bevor empirisch gesicherte Erkenntnisse über soziale und pädagogische Wirkungen vorliegen. Die Linke kritisiert daher: „Union und FDP wollen wahllos in alles investieren, worauf ‚künstliche Intelligenz' steht, ohne zu berücksichtigen, welche ökologischen und sozialen Folgekosten entstehen" (Tatti/Die Linke, BT_251f, 2020, S. 18409). Ferner wird festgestellt, dass die (positive) Wirkung von KI und die mit ihrem Einsatz verbundene Fortschrittsgläubigkeit überbewertet würden (Die Grünen, BT_244b, 2020, S. 8; Janecek/Die Grünen, BT_251f, 2020, S. 18404).

Diese Kritik verweist auf ein Spannungsverhältnis zwischen politischem Handlungsdruck und wissenschaftlicher Evidenz, welches die Legitimation technologiepolitischer Entscheidungen erschwert. Vor diesem Hintergrund sei weitere Forschung zwingend notwendig, die soziale und psychologische Folgen des KI-Einsatzes in den Blick nehme (HFD_1202, 2021, S. 46). Beispielsweise wirft das HFD die staatsbürgerliche Forschungsfrage auf, wie „sich in einer ‚algorithmisierten' Gesellschaft die Mündigkeit von Bürger*innen noch garantieren oder herstellen" (HFD_1202, 2021, S. 10) lasse.

Die Regulierungsdebatte spiegelt diese Spannung wider. Fehlende empirische Erkenntnisse werden dabei mitunter als strukturelles Problem unzureichender Datenzugänge für Wissenschaftler:innen beschrieben (Durz/Union, BT_113d, 2020, S. 23767). Gleichzeitig wird die Notwendigkeit betont, auf Grundlage wissenschaftlicher Erkenntnisse Technikfolgen abschätzen zu können. So gelte es beispielsweise noch zu erforschen, ob Learning Analytics im Vergleich zu Lehrpersonal tatsächlich eine objektivere Leistungsbewertung ermöglicht

(HFD_1210, 2021, S. 2) oder „wie KI-Systeme auf Lernende und Lehrende wirken und wie sie diese dabei unterstützen können, pädagogische Ziele (u. a. Inklusion) zu erreichen“ (BT_113b, 2020, S. 287).

Damit zeigt sich das zentrale Paradox des Collingridge-Dilemmas (siehe Kapitel 1) zwischen der Unsicherheit über Technikfolgen einerseits und dem Bedarf der Technikregulierung andererseits. Diese zeitliche Asymmetrie zwischen Wissen und Steuerbarkeit prägt die politische Praxis und führt zu einer strukturellen Form institutionalisierter Unsicherheit. Für die Cité des Staatsbürgers bedeutet dies eine Gefährdung der Legalitätsnorm, da Regelklarheit unter Bedingungen begrenzter Voraussicht selbst prekär wird. Zugleich führt diese Unsicherheit zu einer Realitätsprüfung des staatsbürgerlichen Anspruchs und zur Forderung nach klaren Regeln und institutioneller Verantwortung. Der nachfolgende Abschnitt befasst sich mit den unterschiedlichen, teils vagen Vorstellungen darüber, wie bestehende Rechtsgrundlagen angepasst und neue regulatorische Strukturen geschaffen werden sollen, um die Legitimität des KI-Einsatzes im Bildungswesen zu sichern.

Diffuse Forderungen nach der Umsetzung bestehender und der Schaffung neuer Rechtsgrundlagen

Im Kontext dieser Unsicherheit wird auf die Vermittlung bereits bestehender Rechtsgrundlagen hingewiesen, durch die KI-einsetzende Hochschulen Bedenken in Hinblick auf Daten- und Persönlichkeitsschutz sowie Urheberrechtsfragen gemindert werden könnten (WR_1177, 2022, S. 10). Dabei steht insbesondere die DSGVO als rechtliche Grundlage im Fokus. Während etwa SPD-Abgeordneter Korkmaz-Emre den europäischen Rechtsrahmen als Ausdruck hoher, international anerkannter ethischer Standards würdigt (BT_113d, 2020, S. 23766), kritisiert die GEW seine unzureichende Wirksamkeit im sensiblen Bildungsbereich, um Persönlichkeitsrechte zu schützen und Überwachung zu verhindern. Sie moniert etwa die Grenzen der Transparenzforderungen an KI, die angesichts selbst für Expert:innen undurchsichtiger algorithmischer Entscheidungsprozesse „ohnehin vollkommen utopisch“ (GEW_1244, 2019, S. 18) seien. Damit wird eine Diskrepanz zwischen formaler Existenz und praktischer Umsetzung von Regelwerken sichtbar. Rechtliche Normen bestehen, ihre Anwendung bleibt jedoch unklar.

Diese Diskrepanz ist zugleich Ausdruck einer reformistischen Kritik, die sich in Forderungen verschiedener wissenschaftlicher und politischer Akteur:innen nach Rechtssicherheit und der Schaffung einheitlicher Standards niederschlägt (Bull-Bischoff/Die Linke, BT_244c, 2019, S. 14627; WR_1177, 2022, S. 89), zum Beispiel in Bezug auf die „Überwachung von Studierenden im Rahmen von Online-Prüfungen sowie den Einsatz von digitalen Technologien (wie z. B. Learning Analytics)“ (WR_1177, 2022, S. 95). Nichtsdestotrotz verbleiben die

Regulierungsvorschläge vor dem Hintergrund der vorherrschenden Unsicherheit vage. Während Expert:innengremien wie das TAB zumindest noch konkrete Regulierungsbedarfe von verpflichtenden Testungen und Auditierungen bis hin zu Einsatzverboten (in Abhängigkeit des Anwendungsfeldes) nennen (TAB_1327, 2021, S. 7), verweisen politische Initiativen auf die diffuse Notwendigkeit „interessenadäquater Transparenz“ (BT_113b, 2020, S. 93), um Persönlichkeitsrechte im Kontext KI-basierter Entscheidungen wahren zu können, oder fordern die Schaffung gemeinsamer Standards für Datennutzung und Weitergabe von Lerndaten, um Vertrauen in KI-Anwendungen im Bildungssystem aufzubauen (BT_288a, 2020, S. 3).

Darüber hinaus gewinnen zunehmend risikobasierte Regulierungsansätze an Bedeutung, wie sie inzwischen im EU AI Act verfolgt werden, in der bildungspolitischen Debatte jedoch noch sehr abstrakt diskutiert werden. Vertreter:innen der Linken und der SPD argumentieren, dass der Grad der Regulierung am Risiko für Gesellschaft und Individuen ausgerichtet werden müsse (Sitte/Die Linke, BT_113d, 2020, S. 23761; Korkmaz-Emre/SPD, BT_113d, 2020, S. 23766). Allerdings bleibt unklar, welche Kriterien eine solche Risikobewertung bestimmen und welche Konsequenzen daraus folgen sollen (TAB_1333, 2021, S. 5). Der Versuch, Unsicherheit durch Regulierung zu kontrollieren, produziert so neue Unsicherheiten über Zuständigkeiten, Bewertungsmaßstäbe und Grenzen politischer Steuerung.

Für die Cité des Staatsbürgers wird diese Unsicherheit über Technikfolgen und Regulierungsbedarfe zur Realitätsprüfung, weil Legalität als moralische Größe an Verbindlichkeit verliert. Darüber hinaus werden auch die vorangegangenen Konflikte zwischen staatsbürgerlichen, solutionistischen, industriellen und marktorientierten Ansprüchen mitunter als Ausdruck einer übergreifenden Regulierungsunsicherheit gelesen. Die Unsicherheit über die Folgen des KI-Einsatzes im Bildungswesen kommt dabei auch im Kontext der Kritik der Cité des Hauses zum Ausdruck, die sich zum Beispiel auf die Gefährdung der Sozialität des Bildungsprozesses oder die Verschiebung tradierter Beziehungen richtet.

9.5.2. Häusliche Kritik der Gefährdung des sozialen Bildungsprozesses durch Künstliche Intelligenz: Vertrauen auf das Tradierte

Die Kritik aus Sicht der Cité des Hauses bildet neben der staatsbürgerlichen Cité den häufigsten moralischen Bezugspunkt zur Hinterfragung des KI-Einsatzes im Bildungswesen. Sie wird in den untersuchten Dokumenten deutlich häufiger zur Kritik als zur Rechtfertigung herangezogen (siehe Abbildung 5), wenngleich sie relational betrachtet einen geringeren Raum in der bildungspolitischen Debatte einnimmt (siehe Abbildung 4). Die Cité des Hauses orientiert sich an Werten wie Stabilität, Vertrauen und Verantwortlichkeit, die aus einem Verständnis sozialer Ordnung und persönlicher Bindung abgeleitet sind. Entsprechend richtet sich ihre Kritik auf

den Erhalt bewährter pädagogischer Praktiken und sozialer Beziehungen, die durch technologische Steuerungslogiken infrage gestellt werden.

Auffällig ist, dass die Kritik der häuslichen Cité in vielen Fällen radikaler Natur ist. Sie zielt nicht nur auf die Ausgestaltung oder Regulierung des Technikeinsatzes, sondern stellt dessen grundsätzliche Sinnhaftigkeit im pädagogischen Kontext infrage. Reformistische Kritik, etwa an einer fehlenden wissenschaftlichen Evidenz oder unzureichenden Rahmenbedingungen, tritt dagegen seltener auf. Im Folgenden werden die verschiedenen Formen häuslicher Kritik systematisch dargestellt. Diese beziehen sich auf fünf zentrale Konfliktfelder: die Abwertung des pädagogischen Selbstverständnisses, die Gefahr der Ersetzung menschlicher Lehrkräfte, die Vernachlässigung der Sozialität des Bildungsprozesses, die Diskrepanz zwischen pädagogischer Bildungslogik und datenbasierter Steuerung sowie die Verschiebung von Kompetenzanforderungen im Zuge des digitalen Wandels. Diese Dimensionen zeigen, wie die Cité des Hauses auf eine moralische Verteidigung pädagogischer Integrität und sozialer Ordnung gegenüber solutionistischen und industriellen Deutungen von Bildung abzielt.

9.5.2.1. Die häusliche Kritik an der Abwertung des pädagogischen Selbstverständnisses im Konflikt mit solutionistischer Steuerung und industrieller Standardisierung

Im Zentrum der häuslichen Kritik steht die Wahrung pädagogischer Werte, die an stabile Rollen, Erfahrungswissen und persönliche Verantwortlichkeit gebunden sind. Sie richtet sich gegen eine technologische Entwicklung, die traditionelle pädagogische Ordnungen infrage stellt und das Selbstverständnis von Lehrkräften verändert. Der KI-Einsatz in Schulen wird dabei nicht nur als funktionale Innovation, sondern als Eingriff in die soziale und moralische Ordnung des Bildungswesens verstanden. So betont die ehemalige Gymnasiallehrerin und Abgeordnete der SPD-Fraktion Völlers, dass der KI-Einsatz dazu führe, dass „die Arbeits- und Lebenswelt Schule verändert werden soll“ (BT_244c, 2019, S. 14626) und weist damit auf eine grundlegende Umgestaltung institutioneller Routinen hin. Diese Veränderung fordere das pädagogische Selbstverständnis der Lehrkräfte heraus, da ihre Rolle als (pädagogische) Fachkräfte zunehmend abgewertet werde (GEW_1314, 2022, S. 2; GEW_1269, 2020, S. 1). In einer zugespitzten Formulierung warnt Espendiller (AfD), Lehrkräfte würden „zu Marionetten von Lernsoftware mit künstlicher Intelligenz degradiert“ (BT_244c, 2019, S. 14548). Diese Kritik verweist auf eine Existenzprüfung, in der die pädagogische Identität der Lehrkraft als moralisch verantwortliche Person infrage gestellt wird, sobald algorithmische Systeme zentrale Aufgaben des Lehrens übernehmen. Vor diesem Hintergrund sei daher auch sicherzustellen, dass Lehrende die Verantwortung im Bildungsprozess behielten (LT_897, 2021, S. 60, 153; HFD_1202, 2021, S. 46).

Aus Sicht der Cité des Hauses richtet sich diese radikale Kritik damit ebenso gegen das Standardisierungsstreben der industriellen Ordnung. Wo industrielle Rationalität Effizienz und Skalierbarkeit priorisiert, betont die häusliche Ordnung Erfahrungswissen, tradiertes Handeln und personale Bindung. Der KI-Einsatz wird als Untergrabung pädagogischer Prinzipien verstanden, weil diese „nicht der Logik der Lerntheorien oder der Entwicklungspsychologie, sondern der Logik der Datenerfassung“ (GEW_1244, 2019, S. 5) folgen. Damit verbunden ist ein Verlust pädagogischer Freiheit (HFD_1191, 2016, S. 21), denn die Lehrendenrolle wandelt sich von Gestalter:in hin zum Mentor:innenstatus in einem durch digitale Systeme vorstrukturierten Lernumfeld (TAB_1335, 2020, S. 203). Damit vollzieht sich eine Existenzprüfung, in der sich Lehrkräfte und ihre Verbände gegen die Vorstellung wenden, dass pädagogische Qualität durch technische Optimierung ersetzt werden könne. Somit steht der Einsatz von KI aus Sicht der Cité des Hauses in einem Konflikt mit solutionistischen Rechtfertigungen des KI-Einsatzes. Der Glaube, dass technologische Systeme Bildungsprobleme lösen könnten, wird als Ausdruck „einer Entwertung menschlichen Erfahrungswissens“ (BT_113b, 2020, S. 309) interpretiert, wodurch das tradierte Selbstverständnis von Arbeit und sozialer Ordnung unter Druck gerät. Die Enquete-Kommission fordert daher eine kritische Auseinandersetzung mit dem Wandel der Arbeitswelt (ebd., S. 309 f.).

Zudem äußert sich der Konflikt mit der Cité der Solution und der Industrie ebenso in Hinblick auf die technokratische Steuerung des Bildungswesens durch Technologieunternehmen. Während im Zuge der staatsbürgerlichen Kritik noch die Sicherung humanistischer Bildungsideale im Vordergrund stand, rückt hier die Souveränität der Lehrkräfte in den Fokus. Bull-Bischoff (Die Linke) betont: „das Pädagogische muss die Federführung behalten, nicht das Konzept von IT-Konzernen und -Lobbyisten“ (BT_244c, 2019, S. 14627). Wenn Bildung durch privatwirtschaftliche Akteur:innen bestimmt wird, verlieren pädagogische Institutionen ihre Handlungsfähigkeit. Zugleich kritisiert die Gegenposition Brandenburgs (FDP), dass die Forderung der stetigen Kontrolle und Steuerung durch Lehrende den Einsatz von KI-Systemen verhindern würde, die eine Vielzahl von Potenzialen besäßen (BT_251f, 2020, S. 18349). Diese diffuse Deutung des Innovationsversprechens verweist auf den Konflikt zwischen pädagogischer Autonomie und technokratischer Steuerung und offenbart somit unterschiedliche Vorstellungen davon, was als legitimer Fortschritt gilt.

9.5.2.2. Die häusliche Kritik des Arbeitsplatzverlustes im Konflikt mit industrieller Rationalisierung und solutionistischem Fortschrittsdenken

Die Gefährdung pädagogischer Autonomie schlägt sich darüber hinaus in der Sorge, dass der Einsatz von KI-Technologien zu einer Reduktion von Arbeitsplätzen führen könnte, nieder

(BT_113b, 2020, S. 555 f.; Röspel/SPD, BT_113d, 2020, S. 23762; HFD_1211, 2021, S. 5). Diese Kritik kann als Ausdruck einer Existenzprüfung im Sinne der Cité des Hauses verstanden werden, da sie die Stabilität sozialer und beruflicher Verhältnisse betrifft. Der drohende Arbeitsplatzverlust wird dabei nicht nur als ökonomisches, sondern als moralisches Problem gesehen, das den Wert menschlicher Arbeit und ihre soziale Einbettung infrage stellt. Damit richtet sich die Kritik zugleich gegen solutionistische und industrielle Vorstellungen technischen Fortschritts. Während die Cité der Industrie die Ersetzung menschlicher Arbeit als Ausdruck von Effizienz und Rationalisierung deutet, erscheint sie aus häuslicher Perspektive als Bedrohung gewachsener sozialer Strukturen und pädagogischer Beziehungskompetenz. Auch die Cité der Solution gerät ins Visier, da die Annahme, technologische Systeme könnten Bildungsprobleme eigenständig lösen, als Ausdruck einer technikzentrierten Logik verstanden wird, die den Menschen aus pädagogischen Prozessen verdrängt. Niedersachsens Kultusminister Tonne (SPD) kritisiert somit eine solutionistische Logik des technikzentrierten Fortschritts:

> All das, was wir hier diskutieren, ist kein Selbstzweck. Wir machen Technik nicht um der Technik willen. Nicht alles, was technisch machbar ist, ist auch vertretbar. Es ist nicht Ziel dieser Landesregierung, Lehrkräfte durch Algorithmen und Lernprogramme zu ersetzen. Bei aller Technik in den Schulen bedarf es immer des Kontakts zwischen Lehrkraft und Schülerinnen und Schülern. (Tonne/SPD, LT-NI_664, 2019, S. 5805)

Diese Position markiert eine klare Grenze zwischen technischer Machbarkeit und pädagogischer Verantwortung. Damit zielt die Kritik zugleich auf eine grundlegende Frage der Machtverteilung im Bildungsprozess: Wer trifft Entscheidungen, und in wessen Verantwortung liegt das Lehren? Bartol (SPD) formuliert diese Abgrenzung, wenn er fordert, zu klären, „welche Entscheidungen wir Maschinen überlassen wollen und welche Entscheidungen zwingend letztlich der Mensch treffen muss“ (BT_113c, 2018, S. 4150). Verschiedene Akteur:innen betonen, dass KI als Unterstützung und nicht als Ersatz dienen müsse (GEW_1285, o. J., S. 2; HFD_1191, 2016, S. 13; Raulfs/SPD, LT-NI_547, 2019, S. 3913; LT-HE_979, 2022, S. 2). Die pädagogische Kernaufgabe, das heißt die Beziehungsarbeit zwischen Lehrenden und Lernenden, könne und dürfe nicht vollständig durch KI ersetzt werden (BT_113b, 2020, S. 297; LT-HE_979, 2022, S. 2; HFD_1191, 2016, S. 13).

Im Spannungsfeld zwischen technologischem Fortschritt und sozialer Verantwortung zeigt sich eine doppelte Kritik. Die häusliche Ordnung wendet sich einerseits gegen die instrumentelle Rationalität der Industrie, die Bildung als Prozess der Optimierung versteht, und andererseits gegen den Solutionismus, der Technologie zum moralischen Leitprinzip erhebt. Die Gefahr des

Arbeitsplatzverlusts wird nicht isoliert betrachtet, sondern als Symptom einer umfassenden Verschiebung der gesellschaftlichen Ordnung verstanden, in der menschliche Arbeit und Verantwortung durch automatisierte Systeme ersetzt werden.

Vor diesem Hintergrund gelte es, die Auswirkungen des KI-Einsatzes auf den Arbeitsmarkt genaustens zu beobachten (Gerber/Die Grünen, LT-SN_904, 2021, S. 2791). Entsprechend müssten sowohl die quantitative Dimension, also der mögliche Wegfall oder die Entstehung neuer Arbeitsplätze, als auch die qualitative Veränderung von Tätigkeiten und Anforderungen berücksichtigt werden (BT_113b, 2020, S. 310). Um negativen Auswirkungen entgegenzuwirken, bedürfe es eines rechtlichen Rahmens, der garantiere, „dass die Kontrolle und Verantwortung immer beim Menschen bleiben“ (Albrecht/Die Grünen, LT-SH_970, 2022, S. 10704). Darüber hinaus sollen Forschung und Lehre einen aktiven Beitrag dazu leisten, „eine funktionierende Mensch-Maschine-Interaktion zu gewährleisten“ (HFD_1202, 2021, S. 48). Diese Forderung knüpft an eine reformistische Realitätsprüfung an, die nicht auf die Abschaffung von KI, sondern auf deren sozial eingebettete Gestaltung und Regulierung abzielt. Sie verbindet technische mit sozialen Anforderungen und betont die Vermittlung zwischen menschlicher und maschineller Handlungskompetenz.

9.5.2.3. Die Aushöhlung der Sozialität des Bildungsprozesses durch solutionistische und industrielle Deutungsmuster

Der Wandel pädagogischer Rollen und die Angst vor der Ersetzung menschlicher Lehrkräfte münden in einen weiteren zentralen Kritikpunkt der häuslichen Cité: die Sorge um die Vernachlässigung der Sozialität des Bildungsprozesses. Bildung wird aus dieser Perspektive nicht als rein kognitiver, sondern als sozialer Vorgang verstanden, der auf persönlicher Interaktion, Beziehung und Vertrauen beruht. Der Einsatz von KI-Technologien wird daher nicht nur als technische Innovation, sondern als Eingriff in die soziale und moralische Ordnung des Bildungswesens wahrgenommen. Befürchtet werden negative Folgen für das Verhältnis zwischen Lehrenden und Lernenden, das als Fundament des pädagogischen Handelns gilt (GEW_1314, 2022, S. 2; HFD_1191, 2016, S. 18). Diese Sorge bezieht sich zugleich auf den gesellschaftlichen Zusammenhalt. Bayaz (Die Grünen) mahnt, es müsse darauf geachtet werden, „dass die Grundsätze unseres Miteinanders nicht ausgehebelt werden“ (BT_113d, 2020, S. 23766), sei Lernerfolg doch „abhängig vom menschlichen Miteinander im sozialen Raum“ (Stumpp/Die Grünen, BT_244c, 2019, S. 14551). Diese Aussagen verweisen auf eine Realitätsprüfung im Sinne der Cité des Hauses: Technologischer Fortschritt wird an sozialen Maßstäben gemessen, insbesondere an der Fähigkeit, zwischenmenschliche Bindungen aufrechtzuerhalten.

In einer radikalen Kritik äußern sich hingegen Vertreter:innen der Linken und der AfD, die die Mensch-Maschine-Interaktion grundsätzlich infrage stellen (Die Linke, BT_288b, 2021, S. 15; Die Linke, BT_406b, 2021, S. 10; Jankowski/AfD, LT-TH_850, 2021, S. 3546). Abgeordnete Bull-Bischoff (Linke) betont: „Bildung ist und bleibt ein sozialer Prozess. Lernen ist und bleibt Beziehungsarbeit zwischen den Lernenden untereinander und zwischen Lernenden und Lehrenden. (…) Soziales Handeln darf nicht auf den Austausch zwischen Mensch und Maschine reduziert werden!“ (BT_244c, 2019, S. 14627). Daher könne Technologie auch nicht die notwendige zwischenmenschliche Interaktion vollständig ersetzen (TAB_1335, 2020, S. 167), denn „der persönliche Kontakt ist nach wie vor und nachweislich sehr entscheidend für den Lernerfolg der Schülerinnen und Schüler“ (LT-BY_993, 2022, S. 12). Andernfalls befürchtet Espendiller (AfD) sogar eine weitere Entfremdung zwischen Lehrenden und Lernenden:

> Unsere Schulen sind heute schon regelrechte Brennpunkte. Schüler, Lehrer und Eltern entfremden sich zunehmend voneinander. Und Sie wollen jetzt unausgereifte Konzepte zur künstlichen Intelligenz in den Unterricht einbauen. Ich kann mir das schon vorstellen: Der Lehrer sagt „Hü!“, die KI sagt „Hott!“. Super, Sie machen damit die Klassenzimmer in diesem Land zum bildungspolitischen Krisengebiet. (BT_244c, 2019, S. 14549)

Diese Aussagen verdeutlichen eine Existenzprüfung: Die soziale und moralische Identität des Bildungswesens wird als gefährdet wahrgenommen, wenn menschliche Nähe durch algorithmische Steuerung ersetzt wird. Auch empirische Beispiele dienen zur Untermauerung dieser Kritik. So berichtet die Soziologin Sigrid Hartong von Eltern im Silicon Valley, die bewusst Privatschulen wählen, um ihren Kindern den Unterricht durch menschliche Lehrkräfte zu sichern (GEW_1244, 2019, S. 15).[35] Diese Beobachtung illustriert, dass soziale Nähe und persönliche Zuwendung als Qualitätsmerkmale von Bildung gelten, die technologische Systeme nicht leisten können. Jankowski (AfD) verweist zudem auf elterliche Bemühungen, Kinder vor einem zu frühen Kontakt mit digitalen Medien zu schützen, und warnt vor deren Untergrabung durch bildungspolitische Digitalisierungsinitiativen (LT-TH_850, 2021, S. 3547).

Auch die GEW sieht in der zunehmenden Digitalisierung eine Gefahr für gewachsene pädagogische Beziehungen. Sie kritisiert die Pläne des Hasso-Plattner-Instituts, mit der sogenannten HPI-Schulcloud Schulen grundlegend zu verändern. Das Konzept sehe vor, „starre, hierarchisches [sic!] Bildungsstrukturen“ (GEW_1260, 2020, S. 2) aufzubrechen und durch den Einsatz

[35] Hartong weist verschärfend darauf hin, dass sich aus diesem Umstand die Frage ergebe, ob „nicht-digitale Bildung (und damit der Entzug von Beobachtung) also langfristig zu einem Privileg der Reichen?“ (GEW_1244, 2019, S. 15) werden würde. Damit wird die Kritik, die hier aus Sicht der Cité des Hauses geäußert wurde, durch eine Kritik der staatsbürgerlichen Cité ergänzt, die eine Verschärfung von Bildungsungleichheiten durch den Einsatz digitaler Technologien anprangert.

von Learning Analytics Lernende zu „Bildungspartnern“ (ebd.) zu machen. Während das Hasso-Plattner-Institut hierin eine Modernisierung sieht, bewertet die GEW diese Entwicklung als Entwertung pädagogischer Prinzipien, da die Leistungsnormierung „dem Primat der Pädagogik“ (ebd.) widerspreche.

In der Gegenüberstellung mit anderen Cités treten die Konfliktlinien klar hervor. Gegenüber der industriellen Ordnung kritisiert die Cité des Hauses insbesondere die funktionale Reduktion des Lernens. Bildung wird hier von der Bundesregierung nicht primär als outputorientierter Prozess verstanden, sondern als zwischenmenschlich geprägte Entwicklung, in der Beziehungsarbeit und individuelle Förderung zentrale Rollen spielen (BT_244b, 2020, S. 9). Es dürfe nicht darum gehen, in möglichst kurzer Zeit möglichst umfassende Lerninhalte erfolgreich zu vermitteln, wenn dies zulasten des gemeinsamen Lernens ginge (Die Linke, BT_244b, 2020, S. 7). Denn das pädagogische Selbstverständnis sieht den Menschen als handelndes, lernendes Subjekt und nicht als Objekt standardisierter Bewertung.

Auch gegenüber der Cité der Solution wird ein Konflikt deutlich. Kritiker:innen betonen, dass Digitalisierungsstrategien oft als Allheilmittel präsentiert werden, während sie in Wirklichkeit pädagogische Konzepte nicht ersetzen können. Es wird moniert, dass allein die Ausstattung mit technischer Hardware fälschlicherweise als Lösung aller Bildungsprobleme dargestellt wird, wobei die „soziale Aufgabe“ (Jankowski/AfD, LT-TH_850, 2021, S. 3546) der Schule vernachlässigt wird. Die Vorstellung, Lehrkräfte könnten durch den KI-Einsatz zu „Marionetten von Lernsoftware“ (Espendiller/AfD, BT_244c, 2019, S. 14548) degradiert werden, wird als bildungspolitische Fehlentwicklung ebenso abgelehnt wie das durch Learning Analytics vermittelte Bild von Lernenden „als irrational, hochgradig emotional beeinflusst und entsprechend beeinflussbar“ (GEW_1244, 2019, S. 10).

Die zunehmende Datenverarbeitung durch KI-Systeme wird in dieser Logik als Symbol einer solutionistischen Ignoranz gegenüber der sozialen Wirklichkeit gedeutet. Das Streben der KI-basierten Messung und Evaluation von Leistungen beabsichtige, „eine abstrakte Lösung für Bildungsprobleme zu entwickeln, anstatt individuelle Förderprogramme zu unterstützen“ (Die Linke, BT_406b, 2021, S. 10). Die GEW kritisiert vor diesem Hintergrund einen KI-basierten Unterricht …

> der tiefgreifende Veränderungen und Eingriffe in Persönlichkeitsrechte der Lernenden, in das Professionsverständnis der Lehrenden und den Umgang zwischen Lernenden und Lehrenden hätte. Dies wirft Fragen auf: Wollen wir KI-gestützten Unterricht? Wie weit

> ist es vom digitalen Unterricht zur digitalen Lehrkraft? Welche Rationalisierungsszenarien stecken in den Forderungen? (GEW_1314, 2022, S. 2)

Mit dieser Position erreicht die häusliche Kritik eine existenzielle Dimension. Sie fechtet nicht mehr nur die Gestaltung, sondern die Legitimität des Technikeinsatzes an, indem erstmalig explizit die Frage aufgeworfen wird, ob KI überhaupt im Bildungssektor eingesetzt werden sollte. Die Cité des Hauses artikuliert somit den moralischen Kern der Auseinandersetzung um Bildung und KI: Sie verteidigt die soziale und zwischenmenschliche Dimension des Lernens gegen solutionistische und industrielle Ordnungen, die den pädagogischen Prozess in eine technologische Logik der Optimierung und Problemlösung überführen wollen.

9.5.2.4. Die Diskrepanz zwischen Pädagogik und Datenerfassung im Konflikt mit industrieller Output-Orientierung

Ein weiterer zentraler Kritikpunkt der häuslichen Cité richtet sich gegen die zunehmende Metrisierung und Quantifizierung von Bildung durch den Einsatz von KI-Technologien. Diese Kritik schließt an die staatsbürgerliche Skepsis gegenüber der Vermessung sozialer Prozesse an, gewinnt hier jedoch eine spezifisch pädagogische Bedeutung. Aus der Perspektive der Cité des Hauses wird weniger die politische Dimension der Datenerhebung problematisiert als vielmehr deren Widerspruch zu pädagogischen Prinzipien des Lehrens und Lernens und damit verbundene negative Konsequenzen für Lernende. Die Einführung datenbasierter Systeme wird als Gefahr für die pädagogische Autonomie und als Ausdruck einer industriellen Logik wahrgenommen, die Bildung nach Kriterien der Effizienz und Messbarkeit organisiert.

Zentral ist dabei die Befürchtung, dass die Metrisierung von Leistung im Widerspruch zu pädagogischen Ansprüchen stehe. In diesem Zusammenhang zitiert die GEW den Verein Digitalcourage, der die Konzentrationsmessung von Schüler:innen mittels EEG-Stirnband und Eyetracking als „Dressur statt Bildung“ (GEW_1269, 2020, S. 1) diskreditiert. Damit richtet sich die Kritik auf den Umgang mit Daten und deren pädagogische Relevanz. So müsse darauf geachtet werden, „dass das Sammeln von Daten kein Selbstzweck wird. Learning Analytics ist pädagogisch erst dann relevant, wenn aus der Unmenge Daten auch wertvolle Aussagen entstehen; denn Lernverhalten ist ein Bildungsprozess, der nicht vollständig statistisch erfasst werden kann“ (Völlers/SPD, BT_244c, 2019, S. 14626). In dieser Aussage zeigt sich eine reformistische Kritik, die den Einsatz datenbasierter Verfahren nicht grundsätzlich ablehnt, ihn aber an pädagogische Sinnhaftigkeit und Kontextsensibilität bindet.

Ilka Hoffmann, Leiterin des Vorstandsbereichs Schule der GEW, hebt die Diskrepanz zwischen pädagogischen Lerntheorien und der Logik der Datenerfassung deutlich hervor:

> Das Neue an Learning Analytics ist, dass die Tools nicht allein von Pädagog*innen entwickelt und kontrolliert werden, sondern durch von Informatiker*innen programmierte Algorithmen. Sie folgen also nicht der Logik der Lerntheorien oder der Entwicklungspsychologie, sondern der Logik der Datenerfassung. Das auch von der Kultusministerkonferenz geforderte Primat der Pädagogik wird hier also außer Kraft gesetzt. Die Fragestellung ist nicht: Wie kann ich die Lernentwicklung besonders gut erfassen und begleiten? Sondern: Welche Daten lassen sich einfach und effektiv erfassen und vergleichen? (GEW_1244, 2019, S. 5)

Diese Aussage verdeutlicht den zentralen Konflikt zwischen häuslicher und industrieller Cité. Während die Pädagogik auf individuelle Förderung und dialogische Entwicklung zielt, orientiert sich die Datenerfassung an Standardisierung, Vergleichbarkeit und Optimierung. Damit einher gehe ein reduktionistischer Ansatz der Leistungsbewertung, der als eine „Dekomplexifizierung" des Bildungsprozesses bezeichnet wird und einen „gravierende[n] Verlust von pädagogischer Beziehungsarbeit" (HFD_1191, 2016, S. 40) zur Folge habe. Gemeint ist, dass Lernprozesse durch algorithmische Systeme in binäre Strukturen von Richtig und Falsch überführt werden, wodurch alternative Lösungswege und soziale Interaktionen an Bedeutung verlieren (GEW_1244, 2019, S. 5; GEW_1314, 2022, S. 2). Indem beispielsweise Learning Analytics den reinen Output stärker als den Lernprozess in den Fokus rücken, führe die Betonung auf Daten und Metriken zu einer Veränderung pädagogischer Bewertungskriterien, die nicht alle Aspekte der Lernentwicklung adäquat erfassen kann (Völlers/SPD, BT_244c, 2019, S. 14626). Diese Outputorientierung wirke sich zudem negativ auf Aufgabenqualität und Unterrichtskonzept aus, da entscheidende Elemente der Lernentwicklung nicht algorithmisch erfasst werden könnten (GEW_1244, 2019, S. 5). Bildungswissenschaftler Deimann betont vor diesem Hintergrund in einem Arbeitspapier des HFD: Wenn der KI-Einsatz dazu führt, dass „nicht mehr ein Lehrender über wichtige pädagogische Fragen entscheidet, sondern Technologien, sollte dies nicht nur als Segen, sondern als eine der zentralen Weichenstellungen diskutiert werden" (HFD_1191, 2016, S. 41).

Die Analyse zeigt somit, dass die Kritik an der Diskrepanz zwischen Pädagogik und Datenerfassung eine doppelte Stoßrichtung besitzt. Einerseits handelt es sich um eine reformistische Kritik, die eine Rückbindung der Technik an pädagogische Prinzipien fordert. Andererseits liegt ihr eine tiefere, existenzielle Dimension zugrunde: die Verteidigung des pädagogischen Selbstverständnisses gegen eine industrielle Logik, die Bildung auf messbare Outputs reduziert. Die häusliche Cité stellt damit eine moralische Gegenbewegung zur Technisierung des

Bildungswesens dar, die den Vorrang der pädagogischen Beziehung gegenüber datengetriebenen Steuerungsmechanismen betont.

9.5.2.5. Die häusliche Kritik des Verlusts tradierter Kompetenzen im Konflikt mit industrieller Rationalisierung und solutionistischer KI-Ausbildung

Neben der Kritik an der Veränderung der Bildungsbeziehung befasst sich die häusliche Cité auch mit den durch den digitalen Wandel entstehenden neuen Kompetenzanforderungen. Während diese vielfach als Notwendigkeit begründet werden (siehe Abschnitte 9.3.1.3, 9.3.2.2 und 9.4.5.2), wird aus häuslicher Perspektive die damit verbundene Veränderung von Rollen und Verantwortlichkeiten kritisch betrachtet. Der Fokus liegt weniger auf der staatsbürgerlichen Forderung nach kritischem Denken als auf der Sorge, dass traditionelle Lern- und Bildungsformen an Bedeutung verlieren und Lernende durch neue Anforderungen überfordert werden könnten.

Im Konflikt mit der industriellen Cité wird deutlich, dass die Digitalisierung des Bildungswesens nicht nur Effizienzgewinne, sondern auch neue Belastungen erzeugt. Die Einführung digitaler Lernformate dient aus industrieller Perspektive der Ressourcenoptimierung und Kostenreduktion, erfordert aber zugleich von den Lernenden ein höheres Maß an Selbststeuerung und Disziplin (TAB_1335, 2020, S. 246). Dies verweist auf eine subtile Verschiebung der Verantwortung im Bildungsprozess, bei der die Entlastung auf struktureller Ebene potenziell mit einer individuellen Überforderung einhergehen kann und damit nicht nur einen Konflikt mit der Cité der Industrie, sondern zugleich eine Realitätsprüfung für die Cité des Hauses darstellt.

Auch gegenüber der Cité der Solution zeigt sich ein Spannungsfeld. Die Forderung, Informatik- und KI-Kompetenzen fest in Lehrplänen zu verankern, steht aus häuslicher Perspektive unter dem Verdacht, traditionelle Inhalte und Kulturtechniken zu verdrängen. So wird auf der einen Seite der Erwerb neuer Kompetenzen von Lehrenden und Lernenden gefordert (HFD_1201, 2021, S. 6; Raulfs/SPD, LT-NI_547, 2019, S. 3913; Baum/FDP, LT-TH_850, 2021, S. 3548; Tischner/CDU, LT-TH_850, 2021, S. 3549; TAB_1335, 2020, S. 28) und betont, Lehrkräfte müssten umfassend aus- und fortgebildet werden, um den veränderten Anforderungen ihres Berufslebens gerecht zu werden (Tischner/CDU, LT-TH_850, 2021, S. 3549; TAB_1335, 2020, S. 167). Auf der anderen Seite wird jedoch die Gefahr einer einseitigen Ausrichtung auf digitale Inhalte kritisiert. So hebt Tiemann (CDU) hervor: „[W]ir brauchen eine sinnvolle Ergänzung der bewährten Inhalte durch digitale Technik, nicht deren Ersatz“ (BT_244c, 2019, S. 14626). Dabei wird die technologische Orientierung als eine Bedrohung kultureller Bildung verstanden:

> Mein nachdrücklicher Wunsch ist es, dass all jene ‚analogen' Fähigkeiten und, ja, Kulturtechniken über der notwendigen Digitalisierung von Bildung nicht aus dem Blick geraten. Für mich ist es in keiner Weise vertretbar, dass meine Bedenken vor einem Verlust ‚traditioneller' Lerninhalte mit einem knappen ‚Das brauchen wir alles nicht mehr' vom Tisch gewischt werden. (Tiemann/Union, BT_244c, 2019, S. 14626)

Diese Position steht im Gegensatz zur solutionistischen Vorstellung, Bildung müsse sich primär an technologischen Zukunftserfordernissen orientieren. Besonders kritisch wird dabei die Frühförderung von KI-Kompetenzen im Grundschulalter gesehen, die von der AfD-Fraktion abgelehnt wird (Jankowski/AfD, LT-TH_850, 2021, S. 3547) und im Widerspruch zum Anspruch alters- und kontextgerechter Bildungspolitik aus der Perspektive der Cité des Hauses steht.

Insgesamt verdeutlicht die Analyse, dass die Kritik der Cité des Hauses am KI-Einsatz im Bildungswesen auf der Verteidigung tradierter pädagogischer Grundprinzipien beruht. Sie richtet sich gegen technologische Rationalisierungen, die Beziehungen, Erfahrungswissen und pädagogische Verantwortung zugunsten von Effizienz oder KI-basierten Problemlösungsansätzen verdrängen. Im Zentrum steht dabei die Sorge um den Verlust pädagogischer Autonomie und um eine Verschiebung der Bildungslogik hin zu technokratischen und ökonomischen Zielvorstellungen. Die Cité des Hauses betont demgegenüber Bildung als Beziehungsarbeit, als sozial eingebetteten und wertorientierten Prozess, in dem menschliche Nähe und Verantwortung zentral bleiben. Damit bildet sie einen moralischen Gegenpol zu industriellen und solutionistischen Logiken und lehnt einen KI-Einsatz mitunter ab, fordert jedoch als Mindestmaß, dass technologische Innovationen den pädagogischen Prinzipien folgen müssen.

9.6. Zusammenfassung der Analyseergebnisse

Für die Analyse der politischen Legitimation des KI-Einsatzes im Bildungswesen wurden drei zentrale Forschungsfragen herausgearbeitet, welche die Untersuchung der moralischen Grundlage der Rechtfertigung (FF1), die Rolle solutionistischer Rechtfertigungen (FF2) sowie die moralische Grundlage der Kritik und resultierende Konfliktlinien (FF3) umfassen. Nachfolgend werden zunächst übergreifende Befunde der Analyse zusammengefasst, bevor die zentralen Erkenntnisse der einzelnen Forschungsfragen im Detail aufgegriffen werden. Dabei offenbart die qualitative Inhaltsanalyse eine bildungspolitische KI-Debatte in Deutschland, die durch eine positive Valenz im Sinne einer flächendeckenden Befürwortung des KI-Einsatzes geprägt ist. KI wird vielfach als Sachzwang dargestellt, der politische Handlungsspielräume vermeintlich vorgibt. Der Einsatz von KI erscheint weitgehend unumgänglich, sodass sich die Debatte vor

allem auf die Frage richtet, wie eine Einführung gelingen kann. Dieser Grundtenor geht mit einem hohen Abstraktionsgrad einher, der sich sowohl in einer diffusen Verwendung des KI-Begriffs als auch in unspezifischen Nutzenversprechen zeigt, denen es häufig an empirischer Evidenz und konkreten Anwendungsbeispielen mangelt. Ähnlich vage bleiben die politischen Vorstellungen über Regulierungsbedarfe des KI-Einsatzes. Gleichzeitig zeigt die Analyse, dass technikoptimistische Annahmen und vielfach unspezifische Nutzenversprechen bereits zur Rechtfertigung umfangreicher Investitionen in die Förderung von KI sowie breit angelegter Initiativen zum Aufbau von KI-Kompetenzen herangezogen werden.

Trotz dieser deterministischen Prägung der bildungspolitischen KI-Debatte wird ein techno-pragmatisches Verständnis deutlich, welches die wechselseitige Beeinflussung von Technik und Gesellschaft anerkennt. So wird eine aktive Gestaltung in Form einer menschenzentrierten KI-Entwicklung gerade deshalb vielfach gefordert, da potenzielle Risiken des KI-Einsatzes, wie zum Beispiel Diskriminierung oder eine Gefährdung von Persönlichkeitsrechten, anerkannt werden. Gleichzeitig scheint der optimistische Tenor angesichts dieser angedeuteten Risiken nicht zu leiden. Stattdessen soll die prospektive Integration ethischer Prinzipien dazu beitragen, die positive Zukunftsvision zu verwirklichen. Dies rückt die moralischen Grundlagen der Rechtfertigung in den Fokus des Erkenntnisinteresses, die diese Vision stützen.

9.6.1. Die moralische Grundlage der politischen Rechtfertigung des Einsatzes Künstlicher Intelligenz

In Hinblick auf die Frage, auf welcher moralischen Grundlage der KI-Einsatz in der politischen Debatte gerechtfertigt wird (FF1), gilt es, nicht nur die relative Häufigkeit der Cités in den analysierten Segmenten zu berücksichtigen, sondern diese im Zusammenhang mit ihrer Valenz und ihrer Verwendung im Rahmen unterschiedlicher Prüfungsformen zu bewerten. Dabei zeigt sich, dass sich die positive Zukunftsvision politischer Akteur:innen vor allem auf die Cités der Industrie und des Marktes stützt. Werden implizit und explizit solutionistische Annahmen als gemeinsame Rechtfertigungslogik gefasst, tritt zudem die Cité der Solution hervor, die im nachfolgenden Abschnitt 9.6.2 berücksichtigt wird. Ergänzend finden sich Rechtfertigungen aus der staatsbürgerlichen Cité, während die Cité des Projekts und die Cité der Meinung nur vereinzelt zur Legitimation des KI-Einsatzes herangezogen werden.

Inhaltlich betonen die verschiedenen Cités unterschiedliche Rechtfertigungsmuster. Die Cité der Industrie hebt Effizienzgewinne, steigende Planungssicherheit und Kontrolle hervor, etwa durch die Möglichkeit, Lernverläufe präziser zu steuern und in der weiteren Konsequenz Ressourcen schonender einzusetzen. Zusätzlich wird der KI-Einsatz durch ein allgemeines Innovationsstreben gerechtfertigt. Die Cité des Marktes rahmt KI vor allem im Hinblick auf globale

Wettbewerbsfähigkeit, weshalb die Förderung von KI als notwendig erscheint, um wirtschaftliche Nachteile zu vermeiden und den Produktionsstandort Deutschland zu sichern. Daneben tritt die Rechtfertigung einer stärkeren Individualisierung des Lernens in den Vordergrund, die als Voraussetzung für eine wettbewerbsfähige Gesellschaft gilt, indem persönliche Fähigkeiten gezielt gefördert und so individuelle Distinktionsmerkmale herausgestellt werden sollen. Die Rechtfertigungen dieser beiden Cités manifestieren sich besonders deutlich in Wahrheitsprüfungen, in denen die Einführung von KI als steuerbar, politisch wünschenswert und notwendig dargestellt wird. Dadurch tragen sie zur Demonstration politischer Handlungsfähigkeit bei, die die gegenwärtige KI-Politik stützt.

Die Cité des Staatsbürgers wird ebenfalls zur Rechtfertigung des KI-Einsatzes herangezogen. In der politischen Debatte zeigt sich dies vor allem in Verweisen auf die Förderung des Gemeinwohls, auf Inklusion sowie auf gesellschaftliche Teilhabe und Mitbestimmung. Gleichzeitig wird argumentiert, dass der KI-Einsatz nicht nur selbst soziale Teilhabe ermöglicht, sondern diese einem sozialkonstruktivistischen Technikverständnis folgend auch voraussetzt, um eine gemeinwohlorientierte KI-Gestaltung sicherzustellen. Die staatsbürgerliche Cité wird somit häufig im Rahmen von Realitätsprüfungen herangezogen, die Herausforderungen einer legitimen KI-Einführung vor dem Hintergrund der Erfüllung moralischer Ansprüche thematisieren. Zudem findet sie in Kompromissen mit den Cités der Industrie und des Marktes Anwendung. So wird beispielsweise der Datenschutz, der aus staatsbürgerlicher Perspektive einen zentralen Schutz von Persönlichkeitsrechten darstellt, in industriellen Kompromissen als vertrauensbildender und innovationsfördernder Rahmen interpretiert. Effizienz- und Innovationsstreben erscheinen dadurch kompatibel mit gemeinwohlorientierten Ansprüchen und werden in einen Kontext von Transparenz und Gerechtigkeit eingebettet. Ähnliche Muster zeigen sich in Kompromissen mit der Cité des Marktes, in denen eine wertebasierte und gemeinwohlorientierte KI-Entwicklung als Alleinstellungsmerkmal gegenüber ausländischen KI-Projekten und damit als potenzieller Wettbewerbsvorteil einer deutschen KI präsentiert wird.

Schließlich werden die Cités des Projekts und der Meinung ebenfalls zur Rechtfertigung des KI-Einsatzes herangezogen, wenngleich deutlich seltener. Die Cité des Projekts rechtfertigt KI als Ausdruck gesellschaftlicher Modernisierung, der die Flexibilisierung des Lernens und Arbeitens unterstützt und starre Hierarchien auflöst. Aus Sicht der Cité der Meinung verspricht der KI-Einsatz vor allem Reputationsgewinne und gesellschaftliche Zustimmung, wodurch KI eine hohe Symbolkraft zugesprochen wird.

9.6.2. Zur Rolle solutionistischer Rechtfertigungen in der politischen Rechtfertigung des Einsatzes künstlicher Intelligenz

In Anbetracht der wachsenden techniksoziologischen Kritik an einem solutionistisch vorangetriebenen KI-Einsatz richtete sich ein weiteres Augenmerk auf die Frage, inwiefern sich solutionistische Rechtfertigungen in der bildungspolitischen Debatte manifestieren (FF2). Die Analyse zeigt, dass solche Rechtfertigungen neben den Bezügen zur Cité der Industrie und des Marktes eine zentrale Grundlage der Befürwortung des KI-Einsatzes bilden. Solutionistische Bezüge treten dabei sowohl explizit als auch implizit auf. Explizite solutionistische Rechtfertigungen beziehen sich auf weitreichende Nutzenversprechen, die KI als Schlüsseltechnologie der Gesellschaft darstellen. Implizite solutionistische Rechtfertigungen äußern sich dagegen in Bereichen, in denen eine politische Rechtfertigung von Maßnahmen wie bereits getätigten Investitionen oder der Forderung der Stärkung von KI-Kompetenzen beispielsweise durch die Einführung des Pflichtfachs Informatik ausbleibt. Der KI-Einsatz erscheint so als selbstverständlich, sodass er keiner weiteren Rechtfertigung mehr zu bedürfen scheint.

Trotz dieser Differenzierung bleiben selbst explizite solutionistische Bezüge inhaltlich vage. Schon die Rechtfertigungen der Effizienzgewinne der Cité der Industrie oder der Wettbewerbsfähigkeit der Cité des Marktes werden selten an konkrete Anwendungen gekoppelt. Für solutionistische Rechtfertigungen gilt dies in besonderem Maße. KI erscheint in einem digitalen Zukunftsgeist als Investition in gesellschaftlichen Fortschritt und wird als technologisches Problemlösungsinstrument mit breiten Potenzialen beschrieben. Diese Potenziale werden in Verweisen auf vermeintlich erfolgreiche Anwendungen und in Auflistungen vielfältiger Einsatzbereiche formuliert, ohne empirisch oder konzeptionell genauer bestimmt zu werden. Häufig wird ein gewissermaßen natürlicher Nutzen für das Bildungswesen unterstellt, wodurch der Einsatz von KI in einem deterministischen Verständnis als alternativlos erscheint.

Solutionistische Rechtfertigungen werden sowohl in funktionelle Kompromisse mit der Cité der Industrie als auch mit der staatsbürgerlichen Cité eingebunden. Im Zusammenspiel mit der Cité der Industrie verbinden sie die Erwartung effizienterer Bildungsprozesse mit der Annahme, dass KI das zentrale Instrument zur Erzielung produktivitätssteigernder Effekte in nahezu allen Anwendungsbereichen sei. In Verbindung mit der staatsbürgerlichen Cité stützen solutionistische Rechtfertigungen beispielsweise die Vorstellung, dass KI die Voraussetzung für mehr Chancengleichheit bilde und dieses Ziel ohne KI nicht erreichbar sei. Gleichzeitig werden solutionistische Ansprüche in Realitätsprüfungen herausgefordert. Politische Akteur:innen verweisen auf die begrenzte Problemlösungsfähigkeit von KI und kritisieren verkürzte Problemdeutungen, die soziale und pädagogische Komplexität nicht berücksichtigen. Zudem wird

beanstandet, dass politische Entscheidungen häufig einem KI-Hype folgen und strukturelle Herausforderungen des Bildungswesens unzureichend adressieren.

9.6.3. Die moralische Grundlage der politischen Kritik des Einsatzes Künstlicher Intelligenz

Diese Erkenntnis knüpft an die Frage nach der moralischen Grundlage der Kritik des KI-Einsatzes und den daraus resultierenden Konfliktlinien an (FF3). Die Kritik beruht vor allem auf den Ansprüchen der staatsbürgerlichen Cité und der Cité des Hauses, die sowohl inhaltliche Einwände formulieren als auch Konflikte mit anderen Cités sichtbar machen. Aus Sicht der staatsbürgerlichen Cité richtet sich die Kritik erstens auf die Gefährdung von Chancen- und Bildungsgerechtigkeit durch mögliche Diskriminierungen und mangelnde Objektivität von KI. Zweitens wird eine Beeinträchtigung der Persönlichkeitsentwicklung im Sinne eines humanistischen Bildungsideals befürchtet, die den Anspruch auf eine kritisch denkende und selbstbestimmte Bürgerschaft unterläuft. Drittens steht die Verletzung von Persönlichkeitsrechten im Vordergrund, insbesondere im Zusammenhang mit Datenschutz. Vereinzelt wird viertens auch eine potenzielle Überwachung durch KI thematisiert, die eine radikale Kritik begründet und den KI-Einsatz infrage stellt, jedoch seltener vertreten wird. Fünftens wird auf die Unsicherheit über die tatsächlichen Folgen des KI-Einsatzes hingewiesen, die sich in einer unklaren Wahrnehmung regulatorischer Anforderungen äußert und durch die Forderung nach weiterer Forschung und zusätzlichen Rechtsgrundlagen reformistisch adressiert wird.

Diese reformistische Kritik, wie sie im Kontext der Cité des Staatsbürgers insbesondere zum Ausdruck kommt, führt zu Realitätsprüfungen, in denen aktuelle Zustände mit normativen Ansprüchen abgeglichen werden. Im Vordergrund stehen die Sicherung eines diskriminierungsfreien KI-Einsatzes sowie der Schutz von Persönlichkeitsrechten im Rahmen einer gemeinwohlorientierten Entwicklung. Zugleich werden durch die Kritik der staatsbürgerlichen Cité Konflikte mit anderen Cités sichtbar, die als Existenzprüfungen verstanden werden. So kollidieren staatsbürgerliche Ansprüche mit solutionistischen Rechtfertigungen, da mit dem KI-Einsatz eine falsche Problempriorisierung kritisiert wird und die technikzentrierte Fokussierung soziale und pädagogische Herausforderungen ausblendet. Ebenso geraten die Effizienz- und Outputorientierung der Cité der Industrie sowie die damit verbundene Metrisierung von Verhalten und Leistung in Konflikt mit humanistischen Bildungsansprüchen und datenschutzbezogenen Bedenken. Darüber hinaus wird auf den Einfluss privatwirtschaftlicher Akteur:innen hingewiesen, der sowohl solutionistische als auch industrielle Logiken stärkt und aus staatsbürgerlicher Sicht eine technokratische Steuerung des Bildungswesens befördert. Schließlich stehen die wettbewerbsorientierten Logiken der Cité des Marktes im Spannungsverhältnis zu einer

gemeinwohlorientierten KI-Entwicklung, während zugleich kritisiert wird, dass datenschutzrechtliche Vorgaben dem Wettbewerb entgegenstehen.

Während die staatsbürgerliche Kritik häufig in Kompromissbildungen mündet, die die Vereinbarkeit von Effizienz, Innovation und Gemeinwohl herstellen sollen, zeigt sich die Kritik politischer Akteur:innen unter Bezug auf die Cité des Hauses deutlich radikaler. Diese tritt zwar seltener auf, mündet jedoch häufiger in Existenzprüfungen, die den KI-Einsatz und die Ansprüche konkurrierender Cités grundsätzlich infrage stellen. Die Cité des Hauses kritisiert erstens die Gefährdung pädagogischer Selbstverständnisse und tradierter Rollen. Daran anknüpfend wird zweitens der Verlust von Arbeitsplätzen befürchtet. Kritisiert wird drittens auch eine Aushöhlung der Sozialität des Bildungsprozesses, die pädagogische Interaktionen und Beziehungsarbeit beeinträchtigen könnte. Hinzu kommen viertens die Diskrepanz zwischen pädagogischen Lerntheorien und datengetriebenen Leistungslogiken sowie fünftens der drohende Verlust tradierter Kompetenzen zugunsten technischer Fähigkeiten. Diese Kritikpunkte richten sich vor allem gegen solutionistische und industrielle Perspektiven, die KI als Effizienzgewinn oder technische Verbesserung darstellen und damit pädagogische Werte und Interaktionen in den Hintergrund rücken. Während die staatsbürgerliche Kritik häufig in funktionale Kompromisse eingebunden wird, bleibt die radikale Kritik der Cité des Hauses in der Regel unbeantwortet. Sie verweist auf richtungsweisende gesellschaftliche Fragen, ob der Einsatz von KI und das damit verbundene Streben nach Effizienz überhaupt ein wünschenswertes Ziel, insbesondere im Bildungswesen, darstellen sollten.

Insgesamt zeigt sich trotz dieser Konfliktlinien eine geringe Intensität konfliktiver politischer Auseinandersetzung. Der überwiegend positive Tenor der Debatte, der sich in einer deutlichen Mehrheit befürwortender Aussagen äußert, geht damit einher, dass die grundsätzliche Frage, ob KI eingesetzt werden sollte, weitgehend ausgeklammert wird. Ebenso werden bestehende Unsicherheiten über die Folgen des KI-Einsatzes weitgehend von technikoptimistischen Erwartungen überlagert. Zwar werden diese Unsicherheit und die daraus resultierenden, häufig abstrakt bleibenden Regulierungsbedarfe vereinzelt zum Gegenstand staatsbürgerlicher Kritik. Zugleich führen sie jedoch nicht dazu, wissenschaftliche Erkenntnisse abzuwarten oder den vorangetriebenen KI-Einsatz zu entschleunigen, der durch Effizienzgewinne und Wohlstandsversprechen attraktiv erscheint.

10. Diskussion

Die vorliegenden Ergebnisse der qualitativen Analyse der Muster der Kritik und Rechtfertigung des KI-Einsatzes in der bildungspolitischen Debatte werden im Folgenden in den konzeptionellen Rahmen dieser Arbeit eingeordnet und im Hinblick auf ihre Konsequenzen für politische Legitimationsprozesse kritisch reflektiert. Ausgehend von der Beobachtung, dass sich der politische Umgang mit KI im Bildungswesen durch eine hohe Diffusität, ein begrenztes Kontingenzbewusstsein und eine niedrige Konfliktintensität auszeichnet, werden im Folgenden drei zentrale Argumentationsstränge entwickelt. Erstens wird erörtert, welche Konsequenzen die Abstraktheit der Debatte für politische Verantwortung und die Gefahr symbolischer Politik hat. Zweitens wird analysiert, wie sich Politisierung im Spannungsfeld von Kontingenz und Konflikt vollzieht und welche Rolle techniksoziologische Deutungsmuster, solutionistische Zukunftsversprechen sowie staatsbürgerliche Rechtfertigungen dabei spielen. Drittens werden daraus die Bedingungen und Grenzen der Legitimation des KI-Einsatzes abgeleitet und mit Blick auf symbolische Stabilisierung, mögliche Legitimitätsbrüche und das technologische Momentum der frühen Gestaltungsphase diskutiert.

10.1. Konsequenzen der Diffusität der politischen Debatte über Künstliche Intelligenz

Die Analyseergebnisse zeigen, dass die bildungspolitische KI-Debatte durch ein hohes Maß an Diffusität gekennzeichnet ist, die sich erstens in der Verwendung des KI-Begriffs, zweitens in der moralischen Rechtfertigung und drittens in den Regulierungsvorschlägen widerspiegelt. Sowohl der KI-Begriff selbst als auch die diskutierten KI-Anwendungen bleiben häufig vage. Dies zeigt sich in der Vermischung digitaler und KI-basierter Lerntechnologien, im seltenen Verweis auf konkrete Anwendungen sowie in diffusen Bezugnahmen auf KI, die weder einem spezifischen Einsatzbereich noch einer klaren Zielsetzung zugeordnet werden können (siehe Abschnitt 9.1).

Neben dem Kontext der KI-Debatte bleiben mitunter auch die moralischen Rechtfertigungen unscharf. Insbesondere Rechtfertigungen der Cité der Industrie und des Marktes reihen abstrakte Nutzenversprechen wie Effizienz, Wohlstand oder Wettbewerbsvorteile aneinander, häufig ohne diese an spezifische Anwendungen zu knüpfen. Ebenso schließt der „digitale Geist“ der solutionistischen Rechtfertigung hier an, der KI als unumgängliches Problemlösungswerkzeug in jedwedem Kontext darstellt. Während die Einsatzbereiche dabei mannigfaltig vorhergesagt werden, vernachlässigt diese mangelnde Kontextualisierung jedoch nicht nur unintendierte Folgen, sondern auch, dass das jeweilige Schadenspotenzial verschiedener KI-Anwendungen erheblich variieren kann (Zweig, 2019; siehe Abschnitt 2.1). Zugleich verlagert sich die

politische Debatte auf eine Metaebene, in der KI zu einem Symbol für Fortschritt und Zukunftsfähigkeit wird, während die konkrete technische oder gesellschaftliche Bedeutung meist unklar bleibt.

Schließlich bleiben auch die Regulierungsvorschläge des KI-Einsatzes vage und gehen in ihrer konkreten Ausgestaltung häufig nicht über die Forderung des Bedarfs selbiger hinaus. Die Diffusität der Debatte verweist somit auf eine strukturelle Unsicherheit im politischen Umgang mit technologischer Komplexität. Diese wird teilweise auch in Forderungen nach wissenschaftlicher Evidenz zur Technikfolgenabschätzung und zu Regulierungsbedarfen adressiert und legt einerseits den Bedarf nach epistemischer Orientierung offen (siehe Abschnitt 9.5.1.4). Andererseits scheinen die abstrakten Nutzenversprechen des KI-Einsatzes diese Unsicherheit jedoch gleichzeitig auszublenden.

Vor diesem Hintergrund werfen die Diffusität und das hohe Abstraktionsniveau der KI-Debatte zwei zentrale Fragen auf, die nachfolgend diskutiert werden: Erstens ist zu fragen, ob politische Akteur:innen einem technologischen Hype verfallen sind, ohne den Gegenstand ihrer Regulierung hinreichend zu verstehen. Zweitens stellt sich die Frage nach der Verantwortlichkeit für die Gestaltung der Technologie.

10.1.1. Symbolpolitik als Konsequenz eines Hypes um Künstliche Intelligenz und als Mechanismus kommunikativer Stabilisierung politischer Handlungsfähigkeit

Die empirischen Erkenntnisse legen nahe, dass die politische Auseinandersetzung mit KI häufig weniger von einem vertieften Verständnis der Technologie als von strategischer Semantik geprägt ist. Die Verwendung des Begriffs KI erfolgt vielfach unscharf und zugleich ausgesprochen positiv konnotiert. Diese positive Aufladung verweist auf ein diskursives Muster, in dem KI als eine Art undefinierter Heilsbringer erscheint (Hirsch-Kreinsen & Krokowski, 2024a; Köstler & Ossewaarde, 2022; Schopmans, 2022) und politische Akteur:innen KI als Symbol gesellschaftlicher Erneuerung aufgreifen, ohne ihre komplexen sozialen Implikationen ausreichend zu verstehen und somit reflektieren zu können. Anstatt klare Kriterien oder Wirkungsnachweise etwa in Form wissenschaftlicher Evidenz für den KI-Einsatz im Bildungswesen heranzuziehen, wird häufig versucht, Legitimität durch die Bezugnahme auf abstrakte Zukunftsversprechen zu erzeugen (z. B. unter Bezug auf die Cité der Industrie oder des Marktes). Dabei wird ein dringender Handlungsbedarf signalisiert, beispielsweise um den Anschluss an die Weltwirtschaft nicht zu verlieren, der auf eine strategische Selbstmobilisierung hindeutet, die das Fehlen konkreter Inhalte durch rhetorische Dynamik kompensiert. In dieser Perspektive erscheint KI-Politik als Ausdruck einer Symbolpolitik, die beabsichtigt, Legitimität performativ zu erzeugen, indem sie auf den wahrgenommenen KI-Hype reagiert und Handlungsfähigkeit

gegenüber der Öffentlichkeit signalisiert, während zugleich jedoch normative und empirische Leerstellen überdeckt werden. Dies zeigt sich etwa in politischer Kommunikation, die den Ausbau von KI-Systemen vorantreibt, obwohl zentrale Fragen der Evidenz oder Technikfolgenabschätzung noch ungeklärt sind. Dabei betonen insbesondere Regierungsvertreter:innen ihre Handlungsfähigkeit in der Kontrolle des technologischen Wandels, wie es insbesondere in Wahrheitsprüfungen der Debatte zum Ausdruck kommt. Diese beabsichtigen, Legitimität nicht durch reflexive Aushandlungsprozesse, sondern durch deren symbolische Bearbeitung und institutionalisierte Selbstvergewisserung herzustellen (Boltanski, 2011).

Die Kritik an einer Symbolpolitik wird in der politischen Debatte selbst explizit aufgegriffen, wenn etwa der FDP vorgeworfen wird, dass diese einem Hype verfallen sei und das Schlagwort KI nutze, ohne eine fundierte inhaltliche Auseinandersetzung mit der Thematik zu leisten (siehe Abschnitt 9.4.4). Während dieser Vorwurf zugleich als Hinweis auf eine allgemeine Unsicherheit über das tatsächliche Verständnis von KI gelesen werden kann, birgt diese Dynamik historisch betrachtet bekannte Risiken. Technologische Hypes stellen Phasen überhöhter Erwartungen dar, die in Enttäuschung und Legitimitätsverlust umschlagen, sobald unerwünschte Folgen sichtbar werden (siehe insbesondere Abschnitt 6.2.2; Brown, 2003; van Lente et al., 2013). Eine Politik, die KI unkritisch als diffuses Heilsversprechen kommuniziert, läuft Gefahr, eine reflektierte Bewertung der Technikfolgen zu behindern, indem Chancen überbetont und Konflikte ausgeblendet werden.

Die in der Diffusität der politischen KI-Debatte sichtbar werdende Symbolpolitik kann dabei als Ausdruck begrenzter staatlicher Kapazitäten und hoher öffentlicher Aufmerksamkeit verstanden werden. Ding (2020) zufolge tritt symbolische Politik[36] insbesondere dann auf, wenn politische Autorität und organisatorische Fähigkeiten gering sind, die gesellschaftliche Beobachtung aber intensiv ist. Unter diesen Bedingungen bemühen sich politische Akteur:innen, durch sichtbare Aktivitäten wie öffentliche Aushandlungsprozesse oder Förderprogramme Responsivität zu demonstrieren und gesellschaftliche Erwartungen zu erfüllen. Die KI-Debatte und die Formulierung von KI-Strategien können damit zum Selbstzweck werden, indem Handlungsfähigkeit demonstriert wird, auch wenn substanzielles Wissen fehlt. Symbolpolitik fungiert so als Instrument, um Vertrauen in politische Handlungsmacht zu sichern.

[36] Ding (2020) verwendet den Begriff der performativen Governance. Im vorliegenden Zusammenhang soll jedoch zwischen symbolischer Politik, die die kommunikative Darstellung politischer Handlungsfähigkeit darstellt, und performativer Politik unterschieden werden. Letztere überführt symbolische Handlungen in tatsächliche Wirksamkeit, etwa wenn politische Zukunftsversprechen durch Investitionen materiell umgesetzt werden und dadurch reale Effekte entfalten (Bareis & Katzenbach, 2022; Borup et al., 2006; siehe vertiefend Abschnitt 10.2.3). Dings Verständnis einer performativen Governance ist an dieser Stelle demnach als Symbolpolitik zu verstehen.

Diese Form der Symbolpolitik kann aus unterschiedlichen Perspektiven betrachtet werden. Erstens wird sie häufig als ineffektiv bewertet, da die mit ihr verbundene Vagheit eine reflektierte politische Steuerung erschwert (Barry & Blühdorn, 2018; Blühdorn, 2007). Bereits im Kontext ethischer Leitlinien der KI-Entwicklung wurde darauf hingewiesen, dass diese aufgrund mangelnder normativer Konkretisierung eine stärkere kommunikative als regulatorische Wirkung entfalten und so in ihrer Effektivität begrenzt bleiben (Hagendorff, 2020; Kerr et al., 2020; siehe Abschnitt 2.4). Zweitens wird Symbolpolitik mitunter als Legitimationsstrategie kritisiert, die vor allem der Stabilisierung politischer Autorität und der Durchsetzung elitärer Interessen dient, indem sie eine kritische Öffentlichkeit beschwichtigt (Edelman, 1976/2005). In dieser Lesart untergräbt sie gesellschaftliches Vertrauen, wenn die verfolgten Interessen dem Gemeinwohl widersprechen (Barry & Blühdorn, 2018). Sichtbar werdende Widersprüche können in einem Scheitern der vertretenen KI-Politik und in öffentlichen Protesten münden. Drittens kann Symbolpolitik eine funktionale Rolle erfüllen, indem sie in komplexen Situationen Orientierung und Zeit verschafft, um Wissen und institutionelle Kapazitäten aufzubauen. Sie legitimiert Autorität, signalisiert moralische Absichten und wird so zu einem zentralen Mittel, um Unsicherheit und Vertrauensfragen kommunikativ zu bewältigen (Blühdorn, 2007). Die Symbolik der diffusen KI-Debatte erscheint damit funktional, da sie politische Unsicherheit bearbeitet und zugleich Vertrauen in die politische Steuerungsfähigkeit sowie in die Kontrolle der KI-Entwicklung zu stabilisieren sucht. Gleichzeitig tritt in den Hintergrund, ob KI-Politik tatsächlich wirksam ist oder differenziert bewertet wird. Diese symbolische Responsivität erfüllt kurzfristig das Bedürfnis nach Orientierung, behindert jedoch langfristig eine reflexive Auseinandersetzung mit den sozialen Folgen der Technologie. So besteht etwa die Gefahr, dass vorschnelle politische Reaktionen einen Deutungsrahmen setzen, der frühe Pfadabhängigkeiten erzeugt, welche die politische Steuerungsfähigkeit langfristig einschränken können (vertiefend hierzu siehe Abschnitt 10.2.3). Ob die politische Debatte über KI Ausdruck einer reflexiven Politik oder einer symbolischen Selbstversicherung ist, kann an dieser Stelle noch nicht abschließend geklärt werden und wird in den nachfolgenden Abschnitten noch einmal aufgegriffen.

Der von Ding (2020) hervorgehobene Kontext der Symbolpolitik einer begrenzten politischen Steuerungsfähigkeit und Autorität verweist zunächst auf eine weiterführende Problematik. Wenn Symbolpolitik Ausdruck einer hohen Diffusität der KI-Debatte ist, droht die Politik, die Verantwortung über die Ausgestaltung und Legitimation der KI zu verlieren.

10.1.2. Mangelnde politische Verantwortung aufgrund des defizitären Verständnisses Künstlicher Intelligenz

Die zuvor herausgearbeitete Ableitung einer praktizierten Symbolpolitik verweist auf ein tieferliegendes Strukturproblem der politischen Auseinandersetzung mit KI in Deutschland. Das defizitär anmutende Verständnis von KI führt dazu, dass politische Akteur:innen ihre Gestaltungskompetenz nicht einlösen können und zunehmend die Fähigkeit verlieren, normative Orientierung in der Entwicklung und Regulierung von KI bereitzustellen. Zugleich eröffnet diese Unschärfe technokratischen und privatwirtschaftlichen Akteur:innen die Möglichkeit, epistemische Autorität über die Deutung und Entwicklung von KI zu beanspruchen. Die aktuelle Technologiepolitik lässt sich dabei als Netzwerkpolitik beschreiben, in der Unternehmen, wissenschaftliche Institutionen und technische Expert:innen ohnehin zentrale Rollen einnehmen (Blümel, 2021; Bogner, 2021b; Heyen et al., 2024; Mai, 2011).

Diese Entwicklung wirft die Frage auf, wer in dieser Konstellation die Verantwortung für eine ethische und soziale Gestaltung von KI trägt. Während staatliche Akteur:innen Legitimation über demokratische Verfahren und moralische Rechtfertigungen im Sinne einer Orientierung am Gemeinwohl herstellen (Boltanski & Thévenot, 1991/2018; Gadinger & Yildiz, 2012), verfolgen privatwirtschaftliche Akteur:innen primär kommerzielle Interessen. Sie legitimieren ihr Handeln über das solutionistische Versprechen, durch KI gesellschaftliche Probleme effizient lösen zu können (Obendiek & Seidel, 2023). Gewinnen diese Versprechen aufgrund eines unzureichenden politischen Verständnisses von KI an Deutungshoheit, verschärfen sich bestehende Machtungleichgewichte zwischen Politik, Wirtschaft und Gesellschaft, wie sie bereits im Zuge der techniksoziologischen Auseinandersetzung diskutiert und auch in der politischen Debatte vereinzelt aufgegriffen wurden (siehe Abschnitte 6.2.1 und 9.5.1.2). Hierdurch gerät die politische Steuerungsfähigkeit unter Druck, da sie sowohl an Einfluss gegenüber dem Markt verliert als auch an demokratischer Rückbindung einbüßt, wenn technologische Entwicklungen zunehmend von nicht gewählten und nicht rechenschaftspflichtigen Akteur:innen geprägt werden. Eine Symbolpolitik überlässt die Gestaltungsmacht jenen, die technologische und ökonomische Ressourcen kontrollieren. Dadurch entsteht ein asymmetrisches Verhältnis, das demokratische Verantwortung schwächt und technokratische Expertise privilegiert, die somit maßgeblich politische Entscheidungen prägt (Bartl, 2023; Pasquale, 2015; Sætra, 2020).

Der wachsende Einfluss großer Technologiekonzerne, der sich nicht ausschließlich auf die KI-Entwicklung beschränkt, sondern zunehmend auch auf ihre Regulierung, wird dabei nicht nur in der wissenschaftlichen Debatte als technokratisches Machtungleichgewicht kritisch bewertet (Benkler, 2019; Hagendorff, 2020; Khanal et al., 2025; Lindman et al., 2023), sondern spiegelt

sich ebenso in Regulierungspräferenzen der Bevölkerung wider, die eine aktive Rolle des Staates in der KI-Regulierung einfordert (König et al., 2023; Wittmann & Meynhardt, 2025; Zhang & Dafoe, 2020). Dabei zeigt sich zum einen im Zeitverlauf eine Zunahme der Befürwortung einer aktiveren Regulierung des Staates und zugleich einer geringen KI-Subventionierung (Marcinkowski & Flaßhoff, 2024). Zum anderen führt ein geringes Vertrauen in Technologieunternehmen dazu, dass nicht bindende Regulierungsmaßnahmen als unzureichend empfunden werden (König et al., 2023). Vor diesem Hintergrund erscheint die Delegation politischer Verantwortung an privatwirtschaftliche und technokratische Akteur:innen fragwürdig. Zwar ermöglicht sie Innovationsdynamiken und entlastet staatliche Institutionen in einem von Unsicherheit geprägten Feld, sie unterläuft jedoch demokratische Steuerungsansprüche und schwächt die moralische Grundlage politischer Legitimität. Im Sinne der SdK bedeutet dies, dass Legitimation nicht mehr aus der Aushandlung unterschiedlicher Cités erwächst, sondern primär durch Effizienz- und Innovationslogiken stabilisiert wird. Damit verschiebt sich die moralische Trägerschaft von Legitimität von einer öffentlichen in eine funktional-technische Ordnung.

Diese Entwicklung verweist auf die strukturelle Begrenztheit politischer Kapazitäten im Sinne der Steuerungsfähigkeit und Autorität (Ding, 2020). Wo Wissen, Autorität und institutionelle Ressourcen fehlen, wird Verantwortung externalisiert. Zugleich ist diese Externalisierung nicht unausweichlich. Politische Akteur:innen verfügen über Handlungsspielräume, in denen sie ihrer gesellschaftlichen Verantwortung gerecht werden können, etwa durch den frühzeitigen Aufbau eigener Expertise oder durch eine proaktive reflexive Ausgestaltung politischer Steuerung. Die Analyse legt jedoch nahe, dass diese Handlungsspielräume bislang nur begrenzt genutzt werden. Dadurch entsteht ein Spannungsverhältnis zwischen symbolischer Handlungsfähigkeit und tatsächlicher Steuerungsverantwortung, dessen Neuausrichtung zentral für die zukünftige politische Gestaltung und Legitimation von KI ist.

10.2. Politisierung Künstlicher Intelligenz im Spannungsfeld von Kontingenz und Konflikt

Politisierung wird in der vorliegenden Arbeit als notwendige Bedingung legitimer Entscheidungsprozesse konzipiert (siehe Kapitel 4), da sie die gesellschaftliche Auseinandersetzung über normative Grundlagen und Gestaltungsoptionen technischer Entwicklungen ermöglicht (Hay, 2007, 2014; Lemke et al., 2024; Meiering & Schäfer, 2020; Schmidt, 2002a; Ulbricht, 2020; Zürn, 2013). Auch wenn der Grad der Politisierung in der vorliegenden Analyse nicht unmittelbar untersucht wurde, erlauben die Ergebnisse zentrale Rückschlüsse auf ihr Ausmaß und ihre Qualität. Im Sinne einer dreidimensionalen Konzeptualisierung lässt sich Politisierung

über die Reichweite der beteiligten Akteur:innen, das Kontingenzbewusstsein politischer Entscheidbarkeit sowie die Konfliktintensität verstehen (de Wilde, 2011; de Wilde et al., 2016; Grande & Kriesi, 2013; Meiering & Schäfer, 2020).

Die Analyse belegt zunächst, dass die *Reichweite der politischen Auseinandersetzung* über KI im Bildungswesen vergleichsweise breit ist. Neben Bundestag und Landtagen beteiligen sich Expert:innengremien, Gewerkschaften und Verbände, wenngleich für den Hochschullehrerbund und den Deutschen Lehrerverband keine öffentlich verfügbaren KI-Dokumente identifiziert werden konnten (siehe Tabelle 5), was auf ein Ungleichgewicht zuungunsten zivilgesellschaftlicher Akteur:innen verweist. Nichtsdestotrotz scheint das Thema fest in der politischen Debatte verankert zu sein. Insofern kann die bildungspolitische Debatte über die KI-Einführung der Minimaldefinition folgend als „Akt des Transports einer Entscheidung (…) in den Bereich des Politischen" (Zürn, 2013, S. 13) als politisiert betrachtet werden. Entscheidend für die Bewertung des Politisierungsgrads sind jedoch weniger die Zahl der beteiligten Akteur:innen als die Art und Tiefe der Auseinandersetzung. Vor diesem Hintergrund richtet sich die folgende Diskussion auf zwei zentrale Dimensionen: Erstens wird das Kontingenzbewusstsein untersucht, das sich in der Anerkennung der politischen Gestaltbarkeit technologischer Entwicklungen zeigt. Dabei werden techniksoziologische Grundannahmen zwischen Determinismus, Konstruktivismus und Pragmatismus herangezogen, um zu diskutieren, inwiefern Politik KI als gestaltbares oder als naturwüchsiges und alternativloses Phänomen begreift. Zweitens wird die Konfliktdynamik betrachtet, die Aufschluss darüber gibt, ob divergierende normative Orientierungen tatsächlich zur Austragung gelangen oder durch Konsensrhetorik überdeckt werden. Darauf aufbauend werden zwei Querschnittsdimensionen vertieft: der Solutionismus als Form kommunikativer Depolitisierung, die Kontingenz und Konflikt systematisch begrenzt, sowie das Potenzial einer integrativen Politisierung, die insbesondere aus staatsbürgerlichen Legitimationsansprüchen hervorgeht. Diese Betrachtungen erlauben zum einen eine strukturierte Diskussion der Ergebnisse unter Rückbezug auf zentrale theoretische Bezugspunkte der Arbeit. Zum anderen ermöglichen sie eine differenzierte Einschätzung, inwiefern die politische KI-Debatte auf eine reflexive Auseinandersetzung mit technologischem Wandel zielt oder primär symbolisch Stabilität herstellt.

10.2.1. Die Einschränkung der Kontingenz als Revival des techniksoziologischen Dualismus?

Die Frage nach dem Kontingenzbewusstsein im politischen Umgang mit KI berührt den Kern politischer Gestaltbarkeit. Dabei wird die Wahrnehmung dieser Gestaltbarkeit wesentlich vom zugrunde liegenden techniksoziologischen Deutungsmuster beeinflusst (siehe Kapitel 6) und betrifft die Einsicht, dass technologische Entwicklungen nicht naturwüchsig verlaufen, sondern

Ergebnis sozialer und politischer Aushandlungsprozesse sind. Die qualitative Untersuchung zeigt jedoch, dass die politische Debatte über KI zwischen technikdeterministischen und sozialkonstruktivistischen Deutungen schwankt, die beide das Ausmaß wahrgenommener Entscheidbarkeit und damit den Grad der Politisierung beeinflussen.

So legt die Analyse offen, dass KI in der bildungspolitischen Debatte wiederholt als zwangsläufige, unumkehrbare Entwicklung beschrieben wird, deren Fortschritt unaufhaltsam und politisch alternativlos erscheint. KI erscheint dabei als Garant für Fortschritt und Effizienz und wird damit, wie Köstler und Ossewaarde (2022) bereits darlegten, zur politisch-administrativen Notwendigkeit (siehe Abschnitt 6.1.2). Bareis und Katzenbach (2022) beschreiben diese Dynamik als diskursive Konstruktion eines dringenden politischen Handlungsbedarfs, welcher durch deterministische Visionen einer verheißungsvollen Zukunft ebenso wie durch Bedrohungsszenarien mobilisiert wird. Politisches Handeln erscheint damit als notwendige Reaktion auf technische Entwicklung. Kontingenz wird zugleich reduziert, weil politische Entscheidungsprozesse durch die im deterministischen Deutungsmuster vorgegebene technische Rationalität übersetzt werden. Eine solche Sichtweise stabilisiert symbolische Politik, da sie Aktivität und Handlungsfähigkeit betont, ohne tatsächliche Gestaltungsmöglichkeiten zu eröffnen. Wenn technologische Entwicklung als eigenlogisch und irreversibel verstanden wird, verliert Politik ihre normative Orientierungsfunktion, und technologische Innovationen treten an die Stelle einer kritischen politischen Gestaltung.

Gleichzeitig verharrt die bildungspolitische KI-Debatte nicht ausschließlich in diesem deterministischen Deutungsmuster, sondern benennt die Unsicherheit über das Verhältnis von Technik und Gesellschaft mitunter explizit. Während dabei auf der einen Seite ein zunehmend dominanter Einfluss der KI auf Bildungsprozesse wahrgenommen wird, wird auf der anderen Seite die aktive Gestaltung der KI-Entwicklung aus einer sozialkonstruktivistischen Perspektive gefordert. Diese Position erfasst partizipative Ansätze als Bedingung und stellenweise auch als Garant einer menschenzentrierten KI-Entwicklung (siehe Abschnitt 9.3.3.2). Teilhabe fungiert hier als Rechtfertigung im Sinne der staatsbürgerlichen Cité, die auf demokratische Kontrolle und Gemeinwohlorientierung abzielt.

Beide Deutungsmuster verlagern den Fokus von der Frage nach der Sinnhaftigkeit des Technikeinsatzes hin zu einer entweder deterministisch-reaktiven oder sozialkonstruktivistisch-aktiven Gestaltung seiner Umsetzung. Die Koexistenz dieser Perspektiven verweist dabei jedoch weniger auf ein Wiederaufleben des klassischen techniksoziologischen Dualismus als auf ein technopragmatisches, wenngleich verengtes Verständnis politischer Kontingenz, das die

wechselseitige Beeinflussung von Technik und Gesellschaft grundsätzlich anerkennt. Dies zeigt sich etwa in der Forderung nach politischer Technikregulierung und der Setzung normativer Standards des KI-Einsatzes, insbesondere mit Blick auf mögliche irreversible Technikfolgen. Auch die Rechtfertigung der Cité des Projekts verweist auf dieses Verständnis, welche KI als Projekt moderner Gesellschaften begreift, das sowohl auf technische Expertise als auch auf politische, wissenschaftliche, wirtschaftliche und gesellschaftliche Mitwirkung angewiesen ist. Gleichzeitig bleibt dieses technopragmatische Verständnis jedoch weitgehend unkritisch. Gestaltungsansprüche richten sich häufig auf die Förderung von Innovation und Wettbewerbsfähigkeit, während Machtverhältnisse und normative Zielkonflikte nur randständig thematisiert werden. Die Debatte erhält dadurch zwar eine konstruktivistische Ausrichtung, verengt sich jedoch auf die Frage, wie die erwarteten Nutzenversprechen bestmöglich realisiert werden können. Das Kontingenzbewusstsein bleibt begrenzt, da politische Gestaltbarkeit zwar anerkannt wird, diese aber im Dienst der technologischen Verwertbarkeit steht, in dem der KI-Einsatz selbst nicht mehr grundsätzlich zur Disposition steht. Eine reflektierte Politisierung, die technologische Unsicherheit als konstitutives Element demokratischer Entscheidbarkeit begreift und die Sinnhaftigkeit der KI-Einführung zum Gegenstand öffentlicher Aushandlung macht, ist bislang nur in Ansätzen erkennbar. So wird bereits deutlich, dass in dieser eingeschränkten Form des Kontingenzbewusstseins eine zentrale Voraussetzung für die geringe Konflikthaftigkeit der KI-Debatte liegt.

10.2.2. Konfliktdynamiken und Prüfungsformen: Zwischen Utopie, Dystopie, Realitätsprüfungen und radikaler Kritik der Cité des Hauses

Die Diskussion der Konfliktdynamik der politischen KI-Debatte gibt Aufschluss darüber, in welchem Maße normative Differenzen sichtbar werden und damit auch, ob Politisierung im Sinne einer reflexiven Aushandlung tatsächlich stattfindet. Die Analyse der Prüfungsformen zeigt, dass über alle Cités hinweg vor allem Wahrheits- und Realitätsprüfungen die Rechtfertigung des KI-Einsatzes in den analysierten Dokumenten dominieren (siehe Abbildung 6), während Existenzprüfungen seltener auftreten. Etwa ein Drittel aller Rechtfertigungen verweist auf Wahrheitsprüfungen, in denen moralische Ansprüche als unstrittig und alternativlos dargestellt werden. Diese Form der Legitimation stabilisiert bestehende Ordnungen, indem sie Kritik ausschließt und die Rechtfertigung des KI-Einsatzes als technisch geboten erscheinen lässt (Boltanski, 2008, 2010; Gadinger & Yildiz, 2012; Smith Ochoa et al., 2021). So werden sowohl Kontingenz als auch Konflikthaftigkeit politischer Entscheidungen reduziert. Besonders ausgeprägt ist diese Alternativlosigkeit im Kontext der Cités der Industrie, des Marktes und der Solution, in denen Wahrheitsprüfungen bis zu 44 % der Rechtfertigungen ausmachen. Die

geringere Häufigkeit von Existenzprüfungen verweist dagegen auf eine begrenzte Austragung moralischer Konflikte. Statt normative Geltungsansprüche offen auszuhandeln, wird Legitimität häufig durch Kompromisse zwischen verschiedenen Cités erzeugt, etwa der des Staatsbürgers, der Industrie, des Marktes und der Solution. Diese pragmatischen Einigungen ermöglichen Übereinkünfte, mindern jedoch die Sichtbarkeit grundlegender normativer Differenzen und begrenzen so den Grad der Politisierung.

Diese Erkenntnis wird durch die überwiegend positive Tonalität der KI-Debatte verstärkt, die in über 80 % der analysierten Segmente zu beobachten ist. Utopische Technikdeutungen prägen mitunter Rechtfertigungen, insbesondere im Falle solutionistischer Bezugnahmen, die KI als universelles Instrument gesellschaftlichen Fortschritts darstellen. Verknüpfen sich utopische mit deterministischen Vorstellungen technologischer Wirksamkeit, erschwert dies eine konfliktive Aushandlung, da technologische Zukunftsvisionen simpel und überzeugend erscheinen und dadurch schwerer hinterfragt werden können (Campolo & Crawford, 2020; Hirsch-Kreinsen & Krokowski, 2024a). Die Diffusität des KI-Begriffs verstärkt diese Wirkung, indem sie konkrete Kritik erschwert und die Debatte auf abstrakte Nutzenversprechen verengt. Dies läuft nicht nur auf eine Schließung von Kontingenz, sondern ebenso auf eine Verkürzung politischer Auseinandersetzung hinaus.

Dystopische Szenarien bilden den Gegenpol zu diesen Zukunftsversprechen. Sie thematisieren vor allem Risiken von Überwachung oder Entmenschlichung, etwa in Bezügen auf Orwell'sche Zukunftsbilder (siehe Abschnitt 9.5.1.3). Entsprechende Szenarien verweisen zwar auf potenzielle Konfliktlinien, bleiben in der Gesamtdebatte aber randständig. Es lässt sich daher weniger von einer Polarisierung als von einem angedeuteten Spannungsfeld zwischen Technikutopie und -dystopie sprechen, welches in Anbetracht der Unsicherheit über die Chancen und Risiken von KI den Raum für abwägende Positionen einengt (Danaher et al., 2017; Elish & boyd, 2018; Nordström, 2022). Gerade dystopische Visionen wirken dabei emotionalisierend und dienen vornehmlich der politischen Abgrenzung und Aufmerksamkeitsgenerierung anstelle einer konstruktiven Aushandlung, wie etwa im Fall der instrumentellen und populistischen Fundamentalkritik der AfD zu beobachten ist (Vorberg, 2019). Dabei instrumentalisieren rechtspopulistische Parteien staatsbürgerliche Kritik in der Bildungspolitik, um diese gegen politische und wissenschaftliche Eliten zu wenden. Indem sie behaupten, dass liberale und wissenschaftliche Akteur:innen ihre eigenen Interessen verfolgen und traditionelle Werte untergraben, stilisieren sie sich als Anwälte des Gemeinwohls und werfen den Eliten vor, Macht zu sichern, statt dem Volk zu dienen (Turnbull et al., 2024). In dieser Logik wird die Cité des Staatsbürgers nicht als

aufrichtige moralische Grundlage genutzt, sondern strategisch instrumentalisiert, um Elitenkritik zu legitimieren und eine konservative Vorstellung des Allgemeinwohls zu behaupten.

Abseits dieser zugespitzten Positionen und rechten Instrumentalisierung zeigt die Analyse dennoch, dass Formen konstruktiver Kritik durchaus vorhanden sind, die im Sinne der SdK die Bezugnahme auf moralische Maßstäbe ermöglichen. Dabei wird (wenngleich keine akteursspezifische Analyse vorgenommen wurde) auch die Rolle der Opposition und der Zivilgesellschaft, in diesem Falle insbesondere der GEW, sichtbar. Aus Perspektive der staatsbürgerlichen Cité werden Fragen demokratischer Teilhabe, Regulierung, Gerechtigkeit und der Wahrung von Persönlichkeitsrechten als zentrale Herausforderungen formuliert. Diese vornehmlich reformistische Kritik zielt jedoch häufig weniger auf grundsätzliche Infragestellung als auf Umsetzbarkeit und gesellschaftliche Verträglichkeit des KI-Einsatzes ab. Damit spielt die Realitätsprüfung eine zentrale Rolle in der gegenwärtigen Debatte. Sie dient dazu, die Vereinbarkeit von Innovation und gesellschaftlicher Ordnung zu prüfen und neu auszutarieren. Im Unterschied zur Wahrheitsprüfung, die Legitimität durch scheinbar objektive Notwendigkeit begründet, bleibt die Realitätsprüfung offen für die Abwägung unterschiedlicher Rechtfertigungsansprüche (Boltanski, 2010; Gadinger, 2016; Smith Ochoa et al., 2021). Dennoch bleibt auch sie in der bildungspolitischen KI-Debatte häufig in einem funktionalen Rahmen verhaftet, indem sie Kritik integriert, ohne eine grundlegende Revision der normativen Grundlagen herbeizuführen.

Insgesamt deutet die Analyse darauf hin, dass die Konfliktdynamik der KI-Debatte durch ein Spannungsverhältnis zwischen Kompromissbildung und selektiver Kritik geprägt ist. Wahrheitsprüfungen und optimistische Zukunftserzählungen stabilisieren bestehende Rechtfertigungsordnungen und legitimieren den KI-Einsatz, während Realitätsprüfungen Kritik in einen moderierenden und funktionalen Rahmen einbetten. Existenzprüfungen, die grundlegende normative Fragen aufwerfen und die Legitimität des KI-Einsatzes selbst infrage stellen, bleiben hingegen selten. Gerade die Cité des Hauses bringt solche radikalen Einwände vor, indem sie auf die Abwertung pädagogischer Professionalität und tradierter Kompetenzen, die Reduktion komplexer Bildungsprozesse auf quantifizierbare Leistungsindikatoren sowie die Aushöhlung der Sozialität des Bildungsprozesses verweist. Diese Kritik greift zentrale Spannungsfelder der techniksoziologischen Auseinandersetzung auf und adressiert nicht primär technische Risiken, sondern zentrale Fragen demokratischer Verantwortung, menschlicher Autonomie und der normativen Grundlagen öffentlicher Bildung. Sie verweist damit darauf, dass demokratische Legitimation nicht allein durch funktionale Problemlösung oder partizipative Verfahren erreicht werden kann, sondern auch die Anerkennung von Grenzen technischer Rationalisierung erfordert. Die geringe Präsenz der Kritik der häuslichen Cité in der Debatte verweist jedoch auf eine

strukturelle Marginalisierung dieser grundlegenden Konflikte. Dies zeigt sich sowohl quantitativ, angesichts des geringen Anteils häuslicher Cité-Bezüge, als auch in ihrer politischen Durchsetzungskraft, indem sie vornehmlich von gewerkschaftlichen Akteur:innen und der Opposition artikuliert wird. Existenzprüfungen finden damit bislang kaum Anschluss an dominante Rechtfertigungszusammenhänge. Zentrale techniksoziologische Problemdiagnosen, etwa zur Verschärfung von Machtungleichgewichten oder zur Verkürzung sozialer Problemlagen durch KI-basierte Quantifizierungslogiken, werden folglich unzureichend in die politische Aushandlung integriert. Damit bleibt ein zentraler Konflikt moderner Technologiepolitik weitgehend ungelöst.

Konflikte werden in der bildungspolitischen KI-Debatte nicht unterdrückt, aber so kanalisiert, dass sie bestehende normative Grundlagen nur begrenzt irritieren, wie insbesondere die Realitätsprüfungen und Kompromisse der Cité des Staatsbürgers verdeutlichen (siehe dazu vertiefend Abschnitt 10.2.4). Eine vertiefte Politisierung, die Konflikt als produktive Ressource demokratischer Legitimität begreift und auch das „Ob“ des KI-Einsatzes verhandelbar hält, ist bislang nur in Ansätzen erkennbar. Im folgenden Abschnitt wird daher diskutiert, inwiefern solutionistische Rechtfertigungen diese Tendenz zur Depolitisierung weiter verstärken.

10.2.3. Die Depolitisierung des Solutionismus: Performative Konsequenzen eines optimistischen Determinismus

Solutionistische Rechtfertigungen prägen den politischen Umgang mit KI in besonderer Weise. Sie äußern sich gewissermaßen als Zuspitzung technokratischer Effizienz- und Wohlstandsversprechen, wie sie etwa im Zuge des Kompromisses solutionistischer und industrieller Rechtfertigungsansprüche zum Ausdruck kommt (siehe Abschnitt 9.4.4) und wie sie bereits in der Technikkritik ab den 1960er Jahren kritisiert wurden (siehe Kapitel 6). Sie verbinden technikdeterministische Grundannahmen mit einem optimistischen Fortschrittsglauben, die sowohl die Kontingenz der politischen Gestaltbarkeit als auch die Konfliktintensität einschränken und somit depolitisierend wirken. Solutionismus lässt technologische Entwicklung als inhärent wünschenswert und politisch notwendig erscheinen. In dieser Perspektive wird KI primär als Mittel zur Lösung gesellschaftlicher, ökonomischer und administrativer Probleme verstanden. Dies mündet in ein Politikverständnis, das gesellschaftliche Herausforderungen zunehmend als technische Steuerungsfragen statt als normative und politische Aushandlungsprozesse begreift. Somit verschiebt sich der Fokus von normativer Reflexion auf technische Machbarkeit, indem politische Debatten auf die Frage reduziert werden, wie KI bestmöglich eingesetzt werden soll, nicht ob ihr Einsatz überhaupt gerechtfertigt ist (Morozov, 2013, 2020; Nachtwey & Seidl, 2021; Obendiek & Seidel, 2023). Die politische Debatte verliert damit ihre kritische Dimension

und droht, zum Schauplatz performativer Handlungsfähigkeit zu werden. Solutionistische Rechtfertigungen verschleiern somit nicht nur Unsicherheiten, sondern erzeugen zugleich den Eindruck rationaler Steuerung in einem komplexen, schwer überschaubaren Feld.

Die empirische Analyse legt nahe, dass das Ausbleiben einer expliziten Rechtfertigung selbst als Ausdruck eines impliziten Solutionismus verstanden werden kann. Politische Forderungen nach dem Ausbau des KI-Einsatzes, technischer Modernisierung oder der Vermittlung von KI-Kompetenzen erfolgen mitunter ohne Rechtfertigung. Dadurch wird der KI-Einsatz nicht als kontingente, sondern als selbstverständlich richtige und notwendige Entwicklung behandelt. Diese implizite Selbstverständlichkeit erscheint problematisch, da sie den Anschein eines gesellschaftlichen Konsenses erweckt, der nicht tatsächlich ausgehandelt wurde. Der Solutionismus fungiert in solchen Fällen als „Default Option" (Morozov, 2020), die eine Rechtfertigung überflüssig erscheinen lässt. Mangelnde Rechtfertigung wird in der vorliegenden Analyse daher als Symptom depolitisierender Selbstverständlichkeit interpretiert, in der die Legitimität technischer Maßnahmen als gegeben vorausgesetzt wird.

Indem implizite wie explizite solutionistische Rechtfertigungen (die in ihrer Gesamtheit in immerhin 30 % der analysierten Segmente nachzuweisen sind) den Fortschritt durch KI als unumgängliche Gegebenheit inszenieren, entsteht ein wachsender politischer Handlungsdruck. Die oftmals einem KI-Hype folgenden Zukunftsvisionen suggerieren eine dringende Notwendigkeit, an den durch KI versprochenen Potenzialen teilzuhaben, sodass die Erwartungen an die Technologie selbst handlungsleitend werden (Borup et al., 2006). Während die im Zusammenhang mit der Diffusität der KI-Debatte diskutierte symbolische Politik diese Zukunftsvisionen vor allem kommunikativ erzeugt (siehe Abschnitt 10.1.1), werden sie im Rahmen solutionistischer Rechtfertigungen zunehmend in die Tat umgesetzt. Sie dienen als Legitimationsgrundlage für politische Maßnahmen wie Investitionen oder die Entwicklung von KI-Kompetenzen. KI-Politik wird damit nicht mehr nur zur Demonstration politischer Handlungsfähigkeit genutzt, sondern erzeugt diese performativ, indem Erwartungen institutionell, insbesondere über finanzielle und strategische Maßnahmen, verankert werden. Die mit diesen Visionen verbundenen politischen Handlungen führen zu Pfadabhängigkeiten, die zukünftige politische Gestaltungsspielräume im Sinne eines Lock-in-Effekts einschränken können (Bareis & Katzenbach, 2022; Borup et al., 2006). So entsteht eine Politik, die einerseits auf Innovationsdruck reagiert, andererseits selbst zur Triebkraft technologischer Entwicklungen wird. Der deterministische Glaube an die Unumgänglichkeit von KI scheint sich dadurch schrittweise in eine selbsterfüllende Prophezeiung zu verwandeln.

Gleichzeitig bleibt fraglich, ob auch der in der politischen Debatte vorherrschende technologische Optimismus das Potenzial besitzt, zur selbsterfüllenden Prophezeiung zu werden. So werden solutionistische Erwartungen politisch verfestigt, bevor empirisch gesicherte Erkenntnisse über soziale Wirkungen der KI vorliegen. Obwohl in verschiedenen Segmenten Unsicherheiten über Technikfolgen und Regulierungsbedarfe zumindest angedeutet werden (siehe Abschnitt 9.5.1.4), wird die politische Förderung von KI weiter vorangetrieben. Dies legt nahe, dass Unsicherheit weniger als Problem erkannt als vielmehr bewusst in Kauf genommen wird. Das Vorgehen verweist einerseits auf das Collingridge-Dilemma (Collingridge, 1980), demzufolge frühe Technikentwicklungen aufgrund mangelnden Wissens schwer regulierbar sind, während in späteren Phasen gewachsene Pfadabhängigkeiten politische Steuerung erschweren. Andererseits birgt eine solche Praxis jedoch Risiken für die Legitimität des KI-Einsatzes, da sie Entscheidungen auf ungesicherten Annahmen gründet und wissenschaftliche Erkenntnisprozesse zugunsten kurzfristiger Handlungsfähigkeit marginalisiert. Solutionistische Rechtfertigungen, aber auch industrie- und marktorientierte Argumente erzeugen einen Handlungsdruck, der an Krisensituationen erinnert. Der KI-Einsatz wird zur dringlichen politischen Aufgabe stilisiert, die keinen Aufschub duldet und kaum Raum für kritische Reflexion oder empirische Validierung lässt.

In diesem Kontext scheint sich die von Bartl (2023) kritisierte Selektivität politischer Wissensverwendung zu bestätigen. Vor dem Hintergrund des wahrgenommenen Handlungsdrucks werden vor allem jene Wissensformen bevorzugt, die schnelle, technisch verwertbare Lösungen anbieten, während komplexe, interdisziplinäre Aushandlungsprozesse, die Ambivalenzen sichtbar machen würden, in den Hintergrund treten (siehe Abschnitt 6.4.4.2). Die politische Praxis droht so einem technokratischen Muster zu folgen, in dem Expertise selektiv eingesetzt wird, um die Wirksamkeit technischer Lösungen zu bestätigen, anstatt sie kritisch zu hinterfragen.

Vor dem Hintergrund der Unsicherheit über die gesellschaftlichen Wirkungen von KI und der gleichzeitig hohen Erwartungen an technologische Lösungen droht die einhergehende Verkürzung politischer und wissenschaftlicher Aushandlungen, den technologischen Optimismus nicht in eine selbsterfüllende, sondern in eine selbstzerstörende Prophezeiung zu verwandeln. Historische Erfahrungen, etwa mit der Entwicklung der Atomkraft (siehe Kapitel 6), zeigen, dass anfängliche positive Technikdeutungen zwar Innovationen und Investitionen mobilisieren, ihre Legitimität jedoch erodiert, wenn die Erwartungen unerfüllt bleiben (Geels & Verhees, 2011). Dies betrifft nicht nur die Technologie selbst, sondern auch die politischen Entscheidungsträger:innen, deren Glaubwürdigkeit leidet, wenn performative Politik Fortschrittsversprechen erzeugt, ohne substantielle Ergebnisse hervorzubringen (Ding, 2020).

Im Kontext der KI-Politik wird somit deutlich, dass sich die symbolische Politik einer diffusen Debatte zunehmend in eine performative Politik verwandelt, die technologische Erwartungen aktiv realisiert. Während dies kurzfristig Handlungsfähigkeit und Legitimität signalisiert, droht langfristig eine Verengung politischer Gestaltungsspielräume. Solutionismus steuert technologische Pfade und reduziert Kontingenz, indem Konflikte und Unsicherheiten überdeckt werden. Damit wird politische Legitimation weniger durch Kritik und Aushandlung als durch performative Fortschrittskommunikation hergestellt. Im folgenden Abschnitt wird daher diskutiert, inwiefern integrative Politisierung diesen Tendenzen entgegenwirken und eine reflexive Legitimation des KI-Einsatzes ermöglichen kann.

10.2.4. Die Cité des Staatsbürgers als heiliger Gral einer integrativen Politisierung Künstlicher Intelligenz

Die bisherige Diskussion des Kontingenzbewusstseins, der Konfliktdynamik und der performativen Politik des Solutionismus verdeutlicht, dass die Politisierung von KI weniger konfliktiv als vielmehr integrativ verläuft. Während offene Auseinandersetzungen tendenziell vermieden werden, zielt die politische KI-Debatte darauf ab, durch die moralische Inklusion staatsbürgerlicher Werte Legitimität herzustellen. Die Cité des Staatsbürgers fungiert dabei als zentrales Bezugssystem, um politische Entscheidungen moralisch abzusichern oder auch zu kritisieren (siehe Abbildung 4). Sie betont Teilhabe, Verantwortung und Gleichheit und bildet damit einerseits einen Gegenpol zu industriellen, marktorientierten und solutionistischen Rechtfertigungslogiken. Andererseits erscheint sie aber auch als eine Art „heiliger Gral“, der gezielt in Kompromisse mit diesen Cités eingebunden wird, um Effizienz und Datenschutz, Gemeinwohl und Wettbewerbsfähigkeit oder technologische Problemlösung und Chancengerechtigkeit zu versöhnen. Die Integration staatsbürgerlicher Ansprüche wird so zu einem zentralen Instrument reflexiver politischer Steuerung, das Legitimität sichern soll, ohne bestehende Macht- und Innovationsstrukturen zu gefährden.

In diesem Spannungsfeld wird die Cité des Staatsbürgers sowohl zur beabsichtigten *Politisierung* als auch zur beabsichtigten *Depolitisierung* herangezogen; wobei beide Strategien letztlich dazu dienen, potenzielle Kontroversen in kontrollierbare Bahnen zu lenken und öffentliche Auseinandersetzungen zu entschärfen. Politisierung zeigt sich dabei in der Betonung von Teilhabe und Mitbestimmung, wodurch eine aktive gesellschaftliche Auseinandersetzung über den Technikeinsatz gefördert und demokratische Kontrolle gestärkt werden soll. Diese Perspektive knüpft an Konzepte partizipativer Governance und Technikfolgenabschätzung an, die seit den 1990er Jahren auf eine Demokratisierung der Technikgestaltung und eine steigende Legitimität technologiepolitischer Entscheidungen durch Partizipation zielen (Bogner, 2021b; Frahm et al.,

2022; Rahm, 2023a). Solche Ansätze entstanden nicht zuletzt als Reaktion auf einen Vertrauensverlust in Expert:innen und politische Entscheidungsträger:innen infolge technischer Fehlschläge (Wehling, 2021; Weyer, 2008). Die gegenwärtige KI-Debatte nimmt sich dieses Versuchs einer prospektiven Legitimitätssicherung durch gesellschaftliche Teilhabe an, die als Voraussetzung einer reflexiven Politisierung verstanden wird. Dadurch sollen normative Orientierungen und gesellschaftliche Zielkonflikte frühzeitig sichtbar und verhandelbar werden (Bogner, 2021b; Doezema & Frahm, 2023; Heyen et al., 2024).

Darüber hinaus legt die Analyse offen, dass sich Depolitisierung dort zeigt, wo politische Akteur:innen staatsbürgerliche Werte wie Gerechtigkeit, Datenschutz und Teilhabe gezielt in die Debatte integrieren, um auf dieser Basis explizit Akzeptanz und Vertrauen in KI zu stärken. Dieses Akzeptanzstreben, das sich auch im Ziel gesellschaftlicher Zustimmung zum KI-Einsatz zeigt, wie es in der Cité der Meinung verankert ist (siehe Abschnitt 9.3.5), kann als strategischer Versuch verstanden werden, die Konfliktintensität der öffentlichen Debatte gering zu halten (Saretzki, 2001; Schimmelfennig, 2021; Zürn, 2013). Politik richtet ihre Kommunikation in dieser Logik darauf aus, über staatsbürgerlich aufgeladene Lösungsrhetorik Zustimmung für bereits vorgezeichnete Pfade zu mobilisieren, anstatt grundlegende Wertkonflikte offen auszutragen (Nokkala, 2016; Ulnicane et al., 2021, 2022). Durch die frühzeitige Einbindung staatsbürgerlicher Ansprüche sollen zukünftige Kontroversen über Regulierung und Technikfolgen präventiv entschärft und die Legitimität politischer Entscheidungen langfristig gesichert werden.

Zugleich birgt dieses Vorgehen das Risiko einer Instrumentalisierung staatsbürgerlicher Werte. Besonders die Kompromisse mit den Cités der Industrie, des Marktes und der Solution zeigen, dass die Cité des Staatsbürgers häufig herangezogen wird, um den KI-Einsatz moralisch abzusichern und potenzielle Kritik abzuschwächen. Ansprüche wie Teilhabe oder Gerechtigkeit können in diesem Kontext von politischen Entscheidungsträger:innen als Mittel zur Durchsetzung des übergeordneten Ziels der KI-Einführung herangezogen werden (siehe insbesondere Abschnitt 9.4.7), obwohl diese Einführung oftmals vor allem durch Erwartungen an Effizienzgewinne oder Wettbewerbsvorteile motiviert ist. Eine solche moralische Vereinnahmung demokratischer Werte droht rhetorische Legitimität zu erzeugen, ohne tatsächliche Steuerungswirkung zu entfalten, wie bereits mit Blick auf die „toothless principles“ der ethischen KI-Entwicklung kritisiert wurde (Kerr et al. 2020; Munn, 2023; Rességuier & Rodrigues 2020; siehe Abschnitt 2.4).

Gleichwohl birgt die Bezugnahme auf staatsbürgerliche Werte ein genuines Potenzial für integrative Politisierung. Dort, wo sie nicht nur als Rechtfertigung, sondern auch als kritischer Bezugspunkt fungiert, eröffnen sich Räume für eine reflektierte Technikgestaltung, in der politische, wirtschaftliche und gesellschaftliche Interessen austariert werden und die wechselseitige Beeinflussung von Technik und Gesellschaft im Sinne eines technopragmatischen Verständnisses anerkannt wird. Wenn es gelingt, stabile Kompromisse zwischen konkurrierenden Cités zu etablieren, etwa zwischen Datenschutz und Innovationsförderung oder zwischen Bildungsgerechtigkeit und Wettbewerbsfähigkeit, kann Legitimation durch Verständigung im Sinne der SdK entstehen.

Die Rolle der Cité des Staatsbürgers in der politischen KI-Debatte bleibt somit ambivalent. Einerseits ermöglicht sie eine integrative Politisierung, die Konflikte konstruktiv einbindet und Legitimität stärkt, andererseits droht sie, durch strategische Instrumentalisierung an normativer Substanz zu verlieren. Eine nachhaltige Legitimation kann daher nur dann gelingen, wenn staatsbürgerliche Prinzipien nicht nur rhetorisch, sondern als handlungsleitende Grundlage politischer Steuerung etabliert werden.

10.3. Konsequenzen für die Legitimation Künstlicher Intelligenz

Die Erkenntnisse der vorliegenden Analyse deuten darauf hin, dass die Legitimation des KI-Einsatzes im Bildungswesen gegenwärtig auf einer fragilen Grundlage steht. So konnte aufgezeigt werden, dass die politische Debatte von drei zentralen Annahmen getragen wird: Erstens erscheint der Einsatz von KI im Sinne einer deterministischen Technikdeutung als unausweichlich. Zweitens dominiert die Vorstellung, dass KI insbesondere vor dem Hintergrund industrieller, marktorientierter und solutionistischer Rechtfertigungen positive Wirkungen für Bildung und Gesellschaft entfalten werde. Drittens wird eine über die Cité des Staatsbürgers legitimierte, menschenzentrierte Ausrichtung als Garant gesehen, antizipierte normative Ansprüche der Gesellschaft an den Technikeinsatz zu erfüllen. Wie die Diskussion der vorangegangenen Abschnitte gezeigt hat, führen diese Annahmen zu einem begrenzten Kontingenzbewusstsein und einer niedrigen Konfliktintensität, die sich in einer insgesamt geringen Politisierung der KI-Debatte widerspiegeln. Sichtbar wird dies darin, dass die Frage der konkreten Ausgestaltung des KI-Einsatzes im Vordergrund steht, während die Frage nach der grundsätzlichen Legitimität der Einführung nur am Rand thematisiert wird. Die politische Auseinandersetzung über KI wird dadurch verkürzt und zentrale Fragen der Legitimation drohen, aus dem Blick zu geraten.

10.3.1. Politisierung als notwendige, nicht jedoch als hinreichende Bedingung legitimer Entscheidungsfindung

Vor diesem Hintergrund stellt sich die Frage, ob Politisierung zwar eine notwendige, jedoch keine hinreichende Bedingung legitimer Entscheidungsfindung darstellt. Die Analyse legt nahe, dass KI durchaus politisiert ist, die moralische Ordnung der Debatte jedoch weitgehend einseitig strukturiert bleibt. Die Legitimation des KI-Einsatzes stützt sich vor allem auf die Cité der Industrie, des Marktes und der Solution, die abstrakte Nutzenversprechen betonen und ihre Geltung häufig durch Wahrheitsprüfungen absichern. Dabei kritisierte bereits Schelsky (1961) die inhärente Legitimität technologischer Effizienzversprechen, solange diese unhinterfragt bleiben (siehe Abschnitt 6.2.1). Durch die Dominanz ökonomischer Rechtfertigungsmuster, wird Kritik nicht systematisch bearbeitet, sondern häufig an den Rand verschoben. Diese Einseitigkeit wirkt sich auf die Verarbeitung von Konflikten aus. Die Cité des Staatsbürgers nimmt dabei eine doppelte Rolle ein. Einerseits artikuliert sie wichtige Einwände, die demokratische Teilhabe, Chancengerechtigkeit und Grundrechte betreffen. Andererseits wird sie funktional in Kompromisse eingebunden, um industriellen und solutionistischen Zielsetzungen moralische Legitimität zu verleihen. Dadurch entsteht eine Form der Konfliktbearbeitung, die zwar integrativ erscheint, kritische Impulse jedoch abschwächt und potenziell konfliktträchtige Fragen in kontrollierbare Bahnen lenkt, indem politische Akteur:innen normative Ansprüche in technische und regulative Detailfragen, etwa über Qualitätsstandards von KI-Systemen, übersetzen.

Dies zeigt sich auch anhand der reformistischen Kritik, die in der politischen KI-Debatte weitgehend produktiv verarbeitet wird. Herausforderungen wie unzureichender Datenschutz oder die fehlende Berücksichtigung pädagogischer Prinzipien werden zwar benannt, jedoch zugleich im Zuge von Realitätsprüfungen als lösbar dargestellt. Wissenschaftliche Evidenz, Regulierung oder technische Anpassungen dienen dabei als Mittel, um bestehende Einwände in die Frage der optimalen Gestaltung des KI-Einsatzes einzubetten. Die grundlegende Problematik der Einführung selbst bleibt so weitgehend unberührt. Radikale Kritik, die die Legitimität der KI-Einführung grundsätzlich infrage stellt, findet hingegen nur selten Gehör. Dies gilt sowohl für staatsbürgerliche Einwände, die auf Risiken für Chancengerechtigkeit und durch Überwachung hinweisen, als auch für häusliche Kritik, die den Verlust sozialer Beziehungen problematisiert. Gerade diese radikale Kritik bleibt häufig unbeantwortet, obwohl sie jene Momente markiert, in denen die Legitimität der KI-Einführung selbst zur politischen Streitfrage wird. Sie zeigt, dass demokratische Politik nicht allein die Gestaltung technischer Entwicklungen bedenken muss, sondern ebenso deren mögliche Begrenzung. Die geringe Resonanz auf solche grundlegenden Einwände verdeutlicht, dass die politische Debatte trotz vorhandener Politisierung

zugunsten industrienaher und solutionistischer Zukunftserzählungen verschoben bleibt. Dies erzeugt eine Spannung zwischen der Notwendigkeit offener Konflikte und der beobachtbaren Tendenz, diese nur selektiv oder symbolisch zu adressieren.

10.3.2. Die fragile Legitimität symbolischer Politik

Vor diesem Hintergrund erscheint die gegenwärtige KI-Politik in zweifacher Hinsicht symbolisch. Zum einen wurden bereits das hohe Abstraktionsniveau und die Diffusität der KI-Debatte als Ausdruck symbolischer Politik kritisch diskutiert (siehe Abschnitt 10.1.1). Zum anderen zeigt sich, dass Konflikte häufig nur oberflächlich oder unzureichend ausgetragen werden, sodass symbolische Politik nicht nur fortbesteht, sondern sich tendenziell verschärft. Die aufgezeigte geringe Politisierung in Verbindung mit einer teilweise symbolisch-performativen KI-Politik kann zu einer fragilen Legitimität des KI-Einsatzes führen. Deutlich wird dies insbesondere, wenn Legitimität nicht nur als Ergebnis eines Aushandlungsprozesses von Rechtfertigung und Kritik im Sinne der SdK verstanden wird, sondern zusätzlich als Wahrnehmung der Angemessenheit politischen Handelns innerhalb gesellschaftlicher Werte und Normen (Suchman, 1995; siehe Abschnitt 3.1). Diese Perspektive stellt keinen Gegensatz zum hier zugrunde gelegten Legitimationsverständnis dar, sondern erweitert es: Während die SdK analysiert, wie Legitimität durch moralische Ordnungen und Aushandlungen erzeugt wird, wird hier die Frage in den Mittelpunkt gerückt, ob diese Rechtfertigungen auch als legitim wahrgenommen werden und damit tatsächlich wirksam werden können. Dabei lassen sich die empirischen Befunde aus einer input- und einer outputorientierten Perspektive diskutieren (vgl. Abschnitt 3.2).

Auf Input-Ebene zeigt sich eine Form symbolischer Politik, die weniger durch substanzielle Problembearbeitung als durch die kommunikative Darstellung politischer Handlungsfähigkeit geprägt ist. Die abstrakte Betonung der Chancen von KI kann kommunikative Input-Legitimität erzeugen, indem sie Responsivität signalisiert und das Gefühl vermittelt, dass politische Akteur:innen gesellschaftliche Erwartungen ernst nehmen (Ding, 2020; Sarcinelli, 1989; siehe Abschnitt 10.1.1). Dies gilt sowohl für die generelle Befassung mit dem KI-Thema als auch für die Bewertung der angeführten Rechtfertigungen, wie beispielsweise im Zuge einer menschenzentrierten KI-Gestaltung.

Wenngleich Legitimation als fortlaufender Prozess gesellschaftlicher Anerkennung zu verstehen ist, der etwa durch neue Ereignisse oder veränderte Technikdeutungen jederzeit erschüttert werden kann (siehe Kapitel 3), verweist diese Form der Legitimation auf ein besonders hohes Maß an Fragilität. Sie setzt voraus, dass politische Entscheidungen tatsächlich wahrgenommene Bedürfnisse der Bevölkerung reflektieren. Genau dies erscheint angesichts vorliegender empirischer Erkenntnisse jedoch zweifelhaft. Untersuchungen zeigen, dass die gesellschaftliche

Themensalienz von KI gering ist und ein niedriges Bewusstsein für ethische Fragen vorherrscht (Kieslich et al., 2023). Gleichzeitig fallen Einstellungen zu KI mitunter ambivalent aus (Brauner et al., 2025; Cave et al., 2019; Kelley et al., 2021; Sindermann et al., 2022) und entwickeln sich im Zeitverlauf eher in Richtung zunehmender Skepsis (Marcinkowski & Flaßhoff, 2024). Gerade im Bildungswesen wird der KI-Einsatz zudem mit einer erhöhten Risikowahrnehmung verknüpft (Lünich et al., 2024a). Vor diesem Hintergrund erscheint es plausibel, dass die politische Förderung von KI weniger auf gesellschaftlichen Erwartungen basiert als auf wirtschaftlichen Interessen. So wurde bereits die Rolle großer Technologieunternehmen in der KI-Entwicklung und -Regulierung sowohl in der wissenschaftlichen Debatte (Khanal et al., 2025; Lindman et al., 2023; siehe insbesondere Abschnitt 6.2.1) als auch vereinzelt in der politischen Debatte kritisch thematisiert (siehe Abschnitt 9.5.1.2).

Auch im Bildungssektor spielen ökonomische Interessen eine zentrale Rolle. In Anbetracht eines schnell wachsenden Marktsegments, das zunehmend von einflussreichen Investor:innen vorangetrieben wird, sind Bildungstechnologien bereits jetzt ein Milliardengeschäft (Williamson, 2024; Williamson & Komljenovic, 2023). Dabei vermarkten große Bildungstechnologieunternehmen positivistisch-deterministische Zukunftserzählungen (Eynon & Young, 2021; Williamson & Komljenovic, 2023), die politische Entscheidungsträger:innen zu beeinflussen scheinen. Williamson (2024) warnt ausdrücklich vor dieser Gefahr:

> Moreover, there is the possibility of political actors buying in to hype over the potential of AI in education, whether generated by academic research centres or commercial edtech companies, and seeking its widespread deployment in schools despite lacking evidence of its intended direct effects (on measurable improvement in learning) or without considering its possible unintended side effects (in a precautionary way). (S. 101)

Was Williamson hier noch als mögliche Gefahr beschreibt, bestätigt die vorliegende Analyse: Die politische KI-Debatte folgt weitgehend einem Hype, der wirtschaftliche und solutionistische Rechtfertigungen privilegiert und damit strukturelle Machtasymmetrien zugunsten politisch-administrativer und unternehmerischer Akteur:innen verstärkt. KI-Politik droht so, zu einem elitengetriebenen Projekt zu werden, das zentrale gesellschaftliche Gruppen von der Mitgestaltung ausschließt und Entscheidungen top-down im vermeintlichen öffentlichen Interesse legitimiert (Hirsch-Kreinsen & Krokowski, 2024b; Köstler & Ossewaarde, 2022). In dieser Perspektive offenbart sich die negative Lesart symbolischer Politik als Stabilisierung politischer Autorität und Durchsetzung elitärer Interessen (Edelman, 1976/2005).

Die vorliegende Analyse kann dabei nicht klären, inwiefern die vorgebrachten Rechtfertigungen in der Gesellschaft tatsächlich als legitim wahrgenommen werden. Doch selbst wenn es der gegenwärtigen KI-Politik gelingt, temporär Input-Legitimität im Sinne gesellschaftlicher Responsivität zu erzeugen, bleibt diese weiter fragil, da sie von der Übereinstimmung zwischen öffentlicher Darstellung und tatsächlicher Wirkung politischer Entscheidungen, also dem Output, abhängt. Die Rechtfertigung des KI-Einsatzes auf Basis von Effizienz, Wettbewerbsvorteilen und Problemlösungsversprechen ist dabei anfällig für einen abrupten Legitimitätsverlust, sobald technologische Anwendungen ihre Versprechen nicht erfüllen oder unerwartete negative Konsequenzen sichtbar werden und folglich ein rekursiver Politisierungsprozess in Gang gesetzt wird (siehe Abschnitt 6.2.2). Die Analyse der KI-Debatte stützt dieses Muster. Industrielle, marktorientierte und solutionistische Rechtfertigungen scheinen glaubwürdig zu wirken, solange Risiken abstrakt sind und kritische Stimmen marginalisiert bleiben. Bleiben Einwände jedoch unbeantwortet, ist eine reaktive Politisierung wahrscheinlich, sobald Zukunftsversprechen enttäuscht werden oder negative Wirkungen des KI-Einsatzes auftreten, wie historische Erfahrungen beispielsweise im Kontext der Atomkraft ebenso aufzeigen wie gegenwärtige Beispiele (König & Wenzelburger, 2021; Ulbricht, 2020; siehe Kapitel 6).

Neben den in dieser Arbeit dargelegten Exempeln des diskriminierenden Notenvergabesystems in Großbritannien sowie dem französischen Studienzulassungsalgorithmus lassen sich entsprechende Entwicklungen auch bereits anhand der Digitalisierung des Bildungsbereichs europäischer Nachbarstaaten beobachten. Dänemark und Schweden galten lange als Vorreiter der digitalen Bildungspolitik, vollzogen jedoch jüngst einen deutlichen Kurswechsel. Der dänische Bildungsminister Mattias Tesfaye entschuldigte sich öffentlich für die flächendeckende Digitalisierung der Schulen mit dem Hinweis: „Unsere Kinder sollten keine Versuchskaninchen in einem digitalen Experiment sein, dessen Ausmaß und Folgen wir nicht absehen können“ (zitiert nach Abraham, 2025, S. 41). Grundlage der Rückabwicklung der Digitalisierung war unter anderem Kritik des schwedischen Karolinska-Instituts, das der Regierung vorwarf, empirische Befunde zu negativen Auswirkungen der Digitalisierung auf Lernleistungen zu ignorieren (Thorell et al., 2023). Das Beispiel macht deutlich, dass eine Vernachlässigung wissenschaftlicher Evidenz und eine Überbetonung visionärer Nutzenversprechen langfristig sowohl technologische als auch politische Legitimität untergraben können. Es zeigt zudem, dass der Prozess reaktiver Politisierung erhebliche politische Verwerfungen nach sich ziehen kann und dass ähnliche Dynamiken auch für KI-Anwendungen antizipiert werden können, sollten diese die hohen Nutzenversprechen nicht einhalten. Dabei wird in der politischen Debatte, vereinzelt im Rahmen der Realitätsprüfung der Cité der Meinung, explizit vor überhöhten Erwartungen und

einem möglichen Vertrauensverlust gewarnt (siehe Abschnitt 9.3.5.2). Gleichzeitig trägt die gegenwärtige KI-Politik durch ihre optimistischen Zukunftsnarrative selbst zur Aufladung dieser Erwartungen bei und verstärkt damit das Risiko späterer Legitimitätskrisen.

10.3.3. Verpasstes Momentum? Politische Legitimation und die Zukunft der Gestaltung Künstlicher Intelligenz

Trotz der historischen Erfahrungen und aktuellen Fallbeispiele droht die gegenwärtige KI-Politik somit, in alte Muster zu verfallen. Techniksoziologische Deutungsmuster und damit verknüpfte Rechtfertigungen moderner Technologien wiederholen sich und der Lernfaktor scheint gering zu sein. In dieser Situation zeichnet sich die Gefahr ab, das technologische Momentum zu verpassen. In den frühen Phasen technologischer Entwicklung besteht noch ein vergleichsweise breiter Handlungsspielraum, in dem gesellschaftliche Anforderungen gestaltbar sind. Wird dieser Zeitraum nicht genutzt, verengen sich die politischen Optionen zusehends und die Technik beeinflusst zunehmend selbst gesellschaftliche Strukturen (Hughes, 1994). Die gegenwärtige performative KI-Politik setzt jedoch bereits Entwicklungen in Gang, die spätere politische Gestaltungsmöglichkeiten einschränken oder mit hohen Kosten verbinden (Bareis & Katzenbach, 2022). Die Analyse zeigt, dass solutionistische Visionen, industrielle Effizienzversprechen und wettbewerbsorientierte Marktlogiken bereits jetzt Pfadabhängigkeiten durch Investitionen und Kompetenzinitiativen erzeugen. Dadurch kann ein zunehmender politischer Druck entstehen, an einmal eingeschlagenen technologischen Pfaden langfristig festzuhalten, weil bereits getätigte Investitionen, entwickelte Infrastrukturen oder politische Zusagen einen Kurswechsel erschweren, selbst wenn Risiken sichtbar werden oder Alternativen verfügbar wären (ein Mechanismus, der als *sunk cost fallacy* bekannt ist). Wie historische Beispiele der Technologiepolitik, etwa der Ausstieg aus der Atomkraft, zeigen, kann ein späterer Richtungswechsel nicht nur finanziell, sondern auch politisch äußerst kostspielig werden und Spannungen zwischen Staat, Institutionen und Gesellschaft verschärfen (Chai et al., 2022; Schiff et al., 2022; von Hirschhausen et al., 2015). Die frühe Engführung der Politisierung auf das „Wie“ trägt dazu bei, dass Alternativen verdrängt werden und die Legitimation des KI-Einsatzes zunehmend davon abhängt, ob die eingeschlagenen Pfade tragfähig sind.

Vor diesem Hintergrund wird deutlich, dass die Legitimation politischer KI-Einführung nicht allein durch funktionale Rechtfertigungen oder durch symbolische Kommunikation hergestellt werden kann. Sie erfordert eine moralisch fundierte, pluralistische und konfliktoffene Auseinandersetzung über Werte, Ziele und Grenzen technischer Entwicklung. Die SdK macht deutlich, dass Legitimität dort entsteht, wo verschiedene moralische Ordnungen in ihrer Spannung zueinander ernsthaft verhandelt werden. Eine solche reflexive Politisierung setzt voraus, dass

das „Ob“ des KI-Einsatzes ebenso zur Disposition steht wie das „Wie“, und dass Konflikte nicht als Störung, sondern als produktive Ressourcen demokratischer Entscheidungsfindung anerkannt werden. Dies gilt insbesondere für die Kritik aus der staatsbürgerlichen und der häuslichen Cité, die auf Gefährdungen demokratischer Kontrolle, sozialer Beziehungen, pädagogischer Praktiken und von Persönlichkeitsrechten verweist. Diese Einwände fungieren als zentrale Prüfsteine, da sie nicht allein technische Risiken problematisieren, sondern grundlegende Voraussetzungen legitimer politischer Ordnung berühren. Die unterschiedlichen Cités stehen damit nicht nur für verschiedene Argumentationsmuster, sondern für konkurrierende Vorstellungen von Fortschritt und gesellschaftlicher Verantwortung im Bildungswesen.

Sollte die Einführung von KI im Bildungswesen tatsächlich im Sinne des Gemeinwohls wünschenswert sein, so kann ihre Legitimation nur gelingen, wenn diese Kritikformen nicht rhetorisch vereinnahmt oder symbolisch beruhigt, sondern aufrichtig und substanziell in die politischen Gestaltungsprozesse integriert werden. Dies betrifft sowohl die staatsbürgerlichen Ansprüche nach Gerechtigkeit und Teilhabe als auch die häusliche Perspektive, die soziale Beziehungen und pädagogische Integrität in den Vordergrund stellt. Eine Politisierung, die diese Perspektiven ernst nimmt, erhöht nicht nur die normative Tragfähigkeit politischer Entscheidungen, sondern eröffnet ebenso Möglichkeiten, technologische Gestaltung bewusst zu begrenzen oder alternative Pfade zu wählen, bevor das technologische Momentum verloren geht.

Die Debatte über KI im Bildungswesen kann daher als exemplarisches Feld gelten, in dem moderne Gesellschaften Legitimität unter Bedingungen technologischer Unsicherheit neu aushandeln. Legitimation kann nur dort gelingen, wo technologische Rationalität und demokratische Verantwortung mindestens gleichberechtigt berücksichtigt werden und wo staatsbürgerliche ebenso wie häusliche Kritik nicht als rhetorische Ressource, sondern als handlungsleitender Bestandteil politischer Steuerung verstanden wird. In diesem Sinne wird die moralische Bearbeitung von Kontingenz und Konflikt selbst zum Prüfstein politischer Legitimität und entscheidet darüber, ob der KI-Einsatz als gesellschaftlich verantwortbar gelten kann oder ob er in künftigen Phasen reaktiver Politisierung scheitern wird.

11. Fazit

Das Ziel dieser Arbeit bestand darin, die Legitimation der politischen Entscheidung über den KI-Einsatz anhand des exemplarischen Anwendungsfalls des Bildungswesens zu untersuchen. Zu diesem Zweck wurde die bildungspolitische Kommunikation auf Grundlage des Legitimationsverständnisses der pragmatischen Soziologie analysiert. Anliegen war es, zu verstehen, wie politische Akteur:innen den Einsatz einer Technologie rechtfertigen, die mit erheblichen

bekannten Risiken und auch unintendierten Folgen verbunden ist. Diese Frage ist angesichts beschleunigter technopolitischer Dynamiken, die sich gegenwärtig durch einen hohen Technikoptimismus auszeichnen, jedoch gleichzeitig mit einem hohen Maß an Unsicherheit über die Wirkungen des KI-Einsatzes einhergehen, von besonderer Bedeutung. Der zunehmend flächendeckende KI-Einsatz scheint nicht selten einem Modus des experimentellen Vorgehens zu folgen, bei dem potenzielle Risiken scheinbar wissentlich in Kauf genommen werden. Ein solcher Umgang nach einem Trial-and-Error-Prinzip berührt grundlegende gesellschaftspolitische Fragen, da nicht nur fehlerhafte Implementierungen, sondern auch unintendierte Wirkungen schwerwiegende Folgen für Lernende, Bildungsinstitutionen und die Gesellschaft haben können. So können etwa Fehlentscheidungen bei KI-basierten Studienzulassungen zu Benachteiligungen führen, die individuelle Bildungswege nachhaltig beeinträchtigen. Ebenso ist denkbar, dass KI-basierte Lernsysteme unbeabsichtigte psychologische oder soziale Folgen hervorrufen, deren langfristige Auswirkungen bislang nicht abzusehen sind. Zugleich bestehen demokratietheoretische Risiken, sollten sich die demokratischen Aushandlungsprozesse über den Einsatz von KI als unzureichend erweisen, sodass Legitimationsdefizite sowohl der KI als auch der politischen Verantwortungszuschreibung drohen.

Vor diesem Hintergrund adressiert die Arbeit ein zentrales Forschungsdesiderat. Während ein großer Teil der bisherigen Forschung die Legitimität, Akzeptanz oder Fairness KI-basierter Entscheidungen untersucht, bleibt die moralische Fundierung der politischen Entscheidung über den Einsatz solcher Systeme weitgehend unbeachtet. Statt einer nachträglichen Bewertung von KI-Entscheidungen ist eine vorgelagerte Betrachtung der politischen Legitimität des KI-Einsatzes erforderlich. Diese ist notwendig, um ethisch fragwürdige Risiken zu vermeiden, demokratische Anforderungen an Entscheidungsprozesse zu erfüllen und dem Verlust öffentlichen Vertrauens vorzubeugen. Dabei wird in der wissenschaftlichen Debatte explizit der Bedarf formuliert, die moralische Grundlage der KI-Politik zum Gegenstand der Forschung zu machen (Boyd & Holton, 2018; Schiff et al., 2022; Tallberg et al., 2023; Züger & Asghari, 2023). Politische Vorstellungen und moralische Ansprüche begleiten technologische Entwicklungen nicht bloß, sondern sie prägen sie aktiv, indem sie Handlungsoptionen strukturieren und gesellschaftliche Veränderungen vorzeichnen (Bareis & Katzenbach, 2022; Perry & Uuk, 2019; Piano, 2020; Taeihagh, 2021). In diesem Kontext gewinnt die Warnung vor einem solutionistisch motivierten KI-Einsatz, der die normative Dimension und Komplexität sozialer Herausforderungen vernachlässigt, zunehmend an Bedeutung. Diese Warnung begründet einen Bedarf, neben den moralischen Grundlagen der KI-Politik auch explizit solutionistische Vorstellungen und ihre Konsequenzen systematisch zu analysieren (Nachtwey & Seidl, 2023). Die Analyse der

politischen Rechtfertigung und Kritik des KI-Einsatzes bildet daher einen wesentlichen Baustein, um die Dynamiken moderner KI-Governance und ihrer Legitimationsprozesse in einem durch Risiko und Unsicherheit geprägten Umfeld zu verstehen.

Zu diesem Zweck entwickelte die Arbeit einen konzeptionellen Rahmen, der die pragmatische SdK, die Politisierungsforschung und techniksoziologische Ansätze der STS miteinander verbindet. Ausgangspunkt ist ein pragmatisches Legitimitätsverständnis, das Legitimation als aktiven Aushandlungsprozess zwischen Kritik und Rechtfertigung begreift. Diese Aushandlung stützt sich auf moralische Ordnungen in Form von Cités, die politischen Entscheidungen Orientierung verleihen und zugleich sichtbar machen, wie unterschiedliche Gemeinwohlbezüge miteinander konkurrieren. Vor diesem Hintergrund versteht die Arbeit Politisierung als notwendige Bedingung legitimer Entscheidungsfindung, da politische Entscheidungen über neue Technologien nur dann legitimiert werden können, wenn sie einem kritischen öffentlichen Aushandlungsprozess zugänglich gemacht und mit moralischen Bezügen gerechtfertigt werden. Ergänzt wird dieser Zugang durch techniksoziologische Deutungsmuster, die zwischen deterministischen, sozialkonstruktivistischen und pragmatistischen Verständnissen technologischen Wandels unterscheiden. Diese Muster prägen sowohl die politischen Rechtfertigungslogiken als auch die Formen der Kritik und ermöglichen es, wiederkehrende historische Konstellationen der Technologiepolitik sichtbar zu machen. Auf diese Weise verbindet der dargelegte Rahmen die Analyse moralischer Ordnungen mit einer Betrachtung politischer Technikdeutungen, die den Umgang mit KI strukturieren. Methodisch wurde dieser konzeptionelle Ansatz in eine inhaltlich strukturierende qualitative Inhaltsanalyse von Policy-Dokumenten zum KI-Einsatz im Bildungswesen übersetzt. Das Vorgehen ermöglichte eine tiefgreifende Rekonstruktion der moralischen Grundlagen und politischen Deutungsmuster, die bildungspolitische Akteur:innen in Rechtfertigungen und Kritiken mobilisieren.

Die empirische Analyse zeigt eine technikoptimistische bildungspolitische KI-Debatte in Deutschland, in der konkrete Vorstellungen von KI sowie die mit ihr verbundenen Nutzenversprechen vage bleiben. Die Legitimation des KI-Einsatzes stützt sich vor allem auf die Cités der Industrie, des Marktes, des Staatsbürgers und der Solution. KI erscheint als unumgängliche Zukunftstechnologie, mit der Effizienzgewinne, Wettbewerbsfähigkeit, Wohlstandsversprechen und gesellschaftlicher Fortschritt in Aussicht gestellt werden. Gleichzeitig werden staatsbürgerliche Ansprüche, etwa an Gemeinwohlorientierung und Gerechtigkeit, häufig in funktionale Kompromisse mit ökonomischen und solutionistischen Rechtfertigungen integriert. Die Debatte ist somit durch ein geringes Kontingenzbewusstsein geprägt: Politische Auseinandersetzungen konzentrieren sich auf die Ausgestaltung der KI-Einführung, während die

Sinnhaftigkeit des Einsatzes ausgeblendet bleibt. Kritik tritt nur randständig auf, etwa in Form von drohender Diskriminierung, Datenschutzproblemen oder einer Gefährdung des humanistischen Bildungsverständnisses. Insbesondere die radikalere Kritik der Cité des Hauses, die auf Spannungen zwischen Standardisierung und pädagogischer Sozialität verweist, bleibt marginal, sodass grundlegende Einwände zur Verschiebung von Verantwortung und Entscheidungsmacht auf technische Systeme kaum Resonanz finden. Die Analyse macht damit sichtbar, wie politische Entscheidungen über den KI-Einsatz gerechtfertigt werden, welche Konfliktlinien sich daraus ergeben und wie Unsicherheit über mögliche Folgen in der politischen Kommunikation bearbeitet wird.

Auf dieser Grundlage trägt die Arbeit erstens zum Verständnis der Legitimation politischer Entscheidungen im Allgemeinen bei, indem sie zeigt, wie politische Akteur:innen Entscheidungen rechtfertigen und welche Rolle moralische Orientierungen dabei spielen. Zweitens eröffnet sie Einblicke in die Bedeutung politischer und normativer Muster von Kritik und Rechtfertigung für die Legitimation moderner Technologien und macht sichtbar, wie techniksoziologische Deutungsmuster diese Prozesse prägen und beeinflussen. Drittens zeigt die Analyse, wie Politik mit komplexen Technologien unter Bedingungen von Unsicherheit umgeht und welche Strategien dabei dominieren. Viertens werden gesellschaftliche und politische Herausforderungen ebenso wie potenzielle Konfliktlinien des digitalen Zeitalters sichtbar, insbesondere im Hinblick auf die Zukunft der Bildung. Schließlich liefert die Arbeit fünftens Lehren für Wissenschaft, Politik und Zivilgesellschaft, etwa in Bezug auf die Sichtbarmachung symbolischer Politik, den Umgang mit technologischer Unsicherheit und die Bedeutung einer reflektierten Auseinandersetzung mit moralischen Rechtfertigungsmustern.

11.1. Lehren der Arbeit für Wissenschaft, Politik und Zivilgesellschaft

Aufbauend auf diesen übergreifenden Beiträgen werden im Folgenden zentrale praktische und reflexive Lehren der Arbeit für Wissenschaft, Politik und Zivilgesellschaft herausgearbeitet. Im Mittelpunkt steht dabei die Frage, welche Einsichten die Untersuchung insbesondere für zukünftige Auseinandersetzungen mit KI und technologischer Transformation besitzt.

11.1.1. Reflexion der wissenschaftlichen Erkenntnisse

Die Arbeit liefert zentrale Erkenntnisse, die zur wissenschaftlichen Auseinandersetzung über politische Legitimationspraktiken, technologische Politikfelder und KI-Governance beitragen. Erstens vertieft die Analyse das Verständnis politischer Legitimation, indem sie zeigt, wie politische Akteur:innen Entscheidungen rechtfertigen und welche Rolle moralische Orientierungen dabei spielen. Die Untersuchung der Cités macht sichtbar, dass Legitimität nicht auf formalen oder technischen Kriterien beruht, sondern in sozialen Aushandlungsprozessen entsteht,

in denen unterschiedliche Gemeinwohlbezüge miteinander konkurrieren und Kritik als notwendiger Bestandteil dieser Prozesse sichtbar wird. Damit wird die Bedeutung einer pragmatischen Perspektive hervorgehoben, die Legitimation als dynamischen, konflikthaften und kontingenten Prozess begreift. Zweitens leistet die Arbeit einen Beitrag zur Weiterentwicklung des Verständnisses von KI-Governance als Politikfeld, welches nicht allein durch technische Regulierungsfragen bestimmt ist, sondern durch moralische Ansprüche, politische Zukunftsnarrative und symbolische Formen der Legitimation. Die Analyse zeigt, dass politische Akteur:innen Unsicherheit häufig symbolisch verarbeiten und Legitimität weniger über empirische Evidenz als über performative Stabilisierung erzeugen. Damit wird deutlich, dass Legitimation technologischer Entscheidungen durch eine politische Praxis geprägt ist, die von Risiken und normativen Spannungen geprägt ist, diese zugleich aber auch verdecken kann.

Darüber hinaus eröffnet die Arbeit drittens einen qualitativen Einblick in politische und moralische Muster der Kritik und Rechtfertigung moderner Technologien, der aus der Verbindung techniksoziologischer Deutungsmuster und der SdK hervorgeht. Diese Kombination erlaubt es, sowohl die diskursive Rahmung technologischer Entwicklungen zu analysieren als auch die moralischen Ordnungen zu erfassen, die das Handeln der politischen Akteur:innen legitimieren sollen. Dadurch wird sichtbar, wie technologische Zukunftsvisionen bestimmte Wertordnungen stärken und politisch plausibel erscheinen lassen, während andere in den Hintergrund treten, und wie sie zugleich politische Handlungsoptionen beeinflussen. Dies erweitert das Verständnis politischer Legitimationsprozesse, indem technologische Zukunftsvisionen und moralische Rechtfertigungspraxen in ihrem Zusammenspiel erfasst werden. Schließlich verdeutlicht der in der Arbeit entwickelte Analyserahmen viertens den Mehrwert einer integrativen Perspektive, die zusätzlich die Rolle der Politisierung miteinbindet. Dieser konzeptionelle Zugang zeigt, wie sich historische Muster der Kritik und Rechtfertigung von Technologien in aktuellen KI-Debatten fortsetzen und wie Politisierung darauf Einfluss nimmt, ob Konflikte sichtbar werden oder verdeckt bleiben. Der Analyserahmen eröffnet damit einen analytischen Zugriff, der sowohl die normative Dimension politischer Entscheidungen als auch die Wechselwirkungen zwischen Technikvorstellungen, Legitimitätsansprüchen und politischer Aushandlung erfasst. Damit kann er für die weitere Forschung politischer Legitimationspraktiken im Allgemeinen sowie technologiepolitischer Entscheidungen im Speziellen herangezogen werden.

11.1.2. Reflexion der Erkenntnisse für die politische Praxis

Für die politische Praxis liefert die Arbeit ebenfalls wesentliche Einsichten, die für eine legitime und reflektierte KI-Politik bedeutsam sind. Die Analyse zeigt erstens, dass politische Entscheidungen über komplexe Technologien unter Bedingungen erheblicher Unsicherheit getroffen

werden und dass diese Unsicherheit häufig kommunikativ überdeckt wird. Die vorliegenden Befunde machen deutlich, dass abstrakte Nutzenversprechen, technikoptimistische Zukunftsbilder und ein weitgehend ungebrochener Fortschrittsglaube die politische Legitimation des KI-Einsatzes prägen, während empirische Evidenzen und kritische Perspektiven nur begrenzt berücksichtigt werden. Insbesondere solutionistische Annahmen lenken die politische Aufmerksamkeit oftmals einseitig auf technologische Instrumente und lassen strukturelle und soziale Herausforderungen des Bildungswesens in den Hintergrund treten. In Anbetracht der Gefahr einer solchen verkürzten Problemwahrnehmung bedarf es daher einer politischen Praxis, die Unsicherheit über die Technikfolgen nicht durch symbolische Stabilisierung überdeckt, sondern sie als integralen Bestandteil technologischer Entscheidungsprozesse offen adressiert. Es bedarf einer reflexiven Aushandlung, die die Ambivalenz technologischer Entwicklung anerkennt und nicht nur die Potenziale, sondern auch die Risiken berücksichtigt, sodass eine ergebnisoffene Gestaltung politischer Entscheidungsprozesse ermöglicht wird. Um der gesellschaftlichen Verantwortung gerecht zu werden, kann der KI-Einsatz dabei nicht im Sinne eines technologischen Innovationsanspruchs als per se gesellschaftsdienlich verstanden werden oder allein an Optimierungserwartungen ausgerichtet werden. Stattdessen muss auch die Möglichkeit einer politischen Begrenzung des KI-Einsatzes Teil des Kontingenzbewusstseins und des politischen Handlungsspielraums sein. Dies erfordert, Kritik nicht nur instrumentell zu verarbeiten, sondern als Ausdruck berechtigter Grundsatzfragen ernst zu nehmen, insbesondere in sensiblen gesellschaftlichen Bereichen wie der Bildung.

Zweitens legt die Arbeit nahe, dass die Legitimation technologiepolitischer Entscheidungen nur dann tragfähig erscheint, wenn Konflikte offen verhandelt und moralische Ansprüche unterschiedlicher Cités nicht zugunsten einer vermeintlich alternativlosen Technologieentwicklung verdrängt oder instrumentalisiert werden. Die Untersuchung zeigt, dass Konfliktlinien zwischen Effizienzlogiken, Gerechtigkeitsansprüchen, pädagogischen Prinzipien und wettbewerbsorientierten Perspektiven bestehen bleiben, auch wenn diese kommunikativ abgeschwächt oder harmonisiert dargestellt werden. Kompromisse, etwa zwischen der staatsbürgerlichen Cité und der Cité der Industrie, können nur legitimieren, wenn sie eine aufrichtige Integration beider Ansprüche ermöglichen und nicht etwa lediglich Effizienzansprüche unter dem Deckmantel einer gemeinwohlorientierten KI vorantreiben.

Drittens weist die Arbeit darauf hin, dass symbolische Politik eine Gefahr für die Legitimation der KI-Politik darstellt. Politische Kommunikation, die sich vor allem auf positive Zukunftsbilder oder nationale Wettbewerbsfähigkeit stützt, riskiert, die Komplexität von KI und die mit ihr verbundenen Risiken zu unterschätzen. Damit drohen Fehler vergangener Technologiepolitik

erneut aufzutreten. Ein technologischer Optimismus sowie die damit verknüpften Rechtfertigungen von Effizienz, Wohlstand und nationaler Wettbewerbsfähigkeit prägen auch heute die politische Debatte und verschärfen sich im Kontext solutionistischer Zukunftsversprechen. Solche Versprechen können jedoch in einer fragilen Legitimität münden, da sie an die Erfüllung hoher Erwartungen geknüpft sind und leicht brüchig werden, wenn unintendierte Technikwirkungen eintreten. Historische Erfahrungen, etwa im Kontext der Atomkraft, zeigen, dass dies zu einer reaktiven Politisierung und zu einem Legitimationsverlust sowohl der Technologie als auch der verantwortlichen politischen Entscheidungsträger:innen führen kann.

Schließlich verdeutlicht die Arbeit die zentrale Rolle der Politik in der Vermittlung zwischen technologischen Möglichkeiten, wirtschaftlichen Interessen und gesellschaftlichen Erwartungen. Eine legitime KI-Gestaltung erfordert die Berücksichtigung unterschiedlicher moralischer Ordnungen, die explizite Einbindung kritischer Perspektiven und den Aufbau belastbarer Evidenzgrundlagen. Eine verstärkte Politisierung, die über symbolische Bezugnahmen hinausgeht und der Kontingenz und Konflikthaftigkeit im Sinne einer expliziten Aushandlung moralischer und gesellschaftlicher Ansprüche Rechnung trägt, stellt daher nicht nur eine notwendige, sondern womöglich auch eine hinreichende Bedingung für demokratisch legitimierte Technologiepolitik dar.

11.1.3. Reflexion der Erkenntnisse aus Sicht der Zivilgesellschaft

Schließlich liefert die Arbeit auch für die Zivilgesellschaft zentrale Einsichten darüber, wie gesellschaftliche Aushandlungen über technologische Innovationen verlaufen und welche Rolle öffentliche Kritik und Teilhabe dabei spielen. Erstens zeigt die Arbeit, dass gesellschaftliche Konfliktlinien im digitalen Zeitalter häufig verdeckt bleiben, wenn technologische Entwicklungen vor allem über positive Zukunftsvisionen und abstrakte Nutzenversprechen legitimiert werden. Die Untersuchung macht Spannungen zwischen Effizienzorientierung, Gemeinwohlansprüchen, der Bedeutung sozialer Interaktionen und wettbewerbsbezogenen Erwartungen sichtbar. Dies ermöglicht ein differenziertes Verständnis dafür, wie KI verschiedene gesellschaftliche Gruppen in unterschiedlicher Weise betrifft und welche normativen Fragen dabei berührt werden. Dabei wird zweitens erkennbar, dass symbolische Politik eine zentrale Herausforderung für das öffentliche Verständnis technologischer Entwicklungen darstellt. Politische Entscheidungen über KI beruhen mitunter auf performativen Bekenntnissen und dem Rekurs auf positive Zukunftsnarrative, die Unsicherheiten und Risiken überdecken. Für die Zivilgesellschaft ergibt sich daraus die Notwendigkeit, politische Kommunikation kritisch zu reflektieren und die dahinterliegenden moralischen und politischen Annahmen zu hinterfragen. Die

Sichtbarmachung symbolischer Politik unterstützt Bürger:innen darin, technologische Versprechen einzuordnen und Macht- sowie Entscheidungsprozesse transparenter zu verstehen.

Drittens verdeutlicht die Analyse, dass die öffentliche Kritik eine zentrale Rolle in der demokratischen Aushandlung technologischer Entwicklungen spielt. Die Perspektiven der staatsbürgerlichen Cité und der Cité des Hauses zeigen, dass insbesondere zivilgesellschaftliche Akteur:innen sensibel für Gefährdungen von Gerechtigkeit, Persönlichkeitsrechten, pädagogischen Prinzipien und sozialen Beziehungen sind. Zugleich sind sie als unmittelbar Betroffene des KI-Einsatzes eine wichtige Stütze einer reflexiven Technikfolgenabschätzung, die Unsicherheit konstruktiv bearbeitet und plurale Perspektiven berücksichtigt. Ihre Kritik trägt dazu bei, problematische Entwicklungen frühzeitig sichtbar zu machen und technologische Entscheidungen nicht ausschließlich technokratischen oder wirtschaftlichen Logiken zu überlassen. Damit erweitern zivilgesellschaftliche Akteur:innen viertens das Bewusstsein für die Kontingenz der KI-Politik, indem sie deutlich machen, dass politischer Handlungsspielraum nicht nur in der Ausgestaltung und Steuerung, sondern ebenso in der Begrenzung des KI-Einsatzes liegen kann. Zugleich rücken sie grundlegendere Fragen über das Verhältnis von Mensch und Technik in den Vordergrund und stellen zur Diskussion, ob ein gesellschaftliches Effizienzstreben, insbesondere im Bildungskontext, überhaupt ein erstrebenswertes Ziel darstellt.

Schließlich macht die Arbeit deutlich, dass der Einbezug kritischer Perspektiven eine wesentliche Voraussetzung für die Legitimation des KI-Einsatzes darstellt und dass zivilgesellschaftliche Akteur:innen eine wichtige Gegenstimme gegenüber einer vorwiegend wirtschaftlichen und politischen KI-Begeisterung bilden. Die Fähigkeit der Gesellschaft, den Wandel des digitalen Zeitalters aktiv mitzugestalten, hängt wesentlich davon ab, wie gut sie die politischen Ziele und moralischen Grundlagen technologiepolitischer Entscheidungen versteht. Die Analyse liefert hierfür Orientierungspunkte, indem sie sichtbar macht, welche Wertvorstellungen, Konfliktlinien und Zukunftsvorstellungen den Einsatz von KI im Bildungswesen prägen.

11.2. Limitationen und anschließende Forschungsbedarfe

Anknüpfend an die Erkenntnisse und Grenzen der vorliegenden Arbeit ergeben sich verschiedene Anschlussfragen für die weitere Forschung. Dabei ermöglichte die qualitative Herangehensweise eine detaillierte inhaltliche Rekonstruktion der moralischen Grundlagen der bildungspolitischen KI-Debatte und lieferte erste Hinweise darauf, dass etwa Gewerkschaften eher kritische und Regierungsparteien überwiegend befürwortende Positionen vertreten. Zugleich erlaubt dieser qualitative Zugang nur begrenzte Aussagen über die quantitative Verteilung akteursspezifischer Rechtfertigungsmuster innerhalb der Gesamtdebatte. Vor dem Hintergrund

der dargelegten Differenzen erscheint daher eine vertiefende Analyse sinnvoll, die akteursspezifische Muster von Kritik und Rechtfertigung systematisch in den Blick nimmt und die Sichtbarkeit einzelner Positionen im Verhältnis zur Gesamtdebatte untersucht. So könnten etwa Erkenntnisse darüber gewonnen werden, ob kritische Perspektiven strukturell unterrepräsentiert sind oder ob befürwortende Positionen durch institutionelle Machtasymmetrien verstärkt werden.

Hinzukommt die Frage nach der Konstruktion politischer Standpunkte. Die Analyse offenbart eine Dominanz wirtschaftlicher und solutionistischer Rechtfertigungen, die auf kommerzielle Interessen verweisen und Technologieunternehmen zugeschrieben werden. Vor dem Hintergrund der Kritik an wirtschaftlicher Einflussnahme und der politischen Übernahme technikzentrierter Zukunftsvisionen stellt sich daher die Frage, inwiefern privatwirtschaftliche Interessen politische Vorstellungen über Chancen und Risiken von KI und damit politische Kritik und Rechtfertigung prägen. Untersuchungen etwa zu Lobbykontakten könnten aufzeigen, ob direkte Interaktionen mit Unternehmen akteursspezifische Zukunftsnarrative begünstigen.

Eine weitere Limitation ergibt sich aus der aggregierten Rekonstruktion der Debatte über den KI-Einsatz im Bildungswesen. Dies ermöglicht es zwar, normative Konflikte auch dann sichtbar zu machen, wenn sie nicht unmittelbar in der direkten Konfrontation verschiedener Rechtfertigungsansprüche unterschiedlicher Akteur:innen auftreten. Ein stärker sequenzorientierter und dialogischer Ansatz, der einzelne Policy-Entscheidungen und ihre argumentative Entwicklung untersucht, könnte hingegen aufzeigen, wie Rechtfertigungen und Kritik im Prozessverlauf unmittelbar aufeinander reagieren. Ebenso könnte er zur Klärung beitragen, ob der in dieser Arbeit dargelegte implizite Solutionismus im Sinne einer mangelnden Rechtfertigung tatsächlich Ausdruck eines deterministischen Optimismus ist, der den KI-Einsatz als selbstverständlich erscheinen lässt, oder ob fehlende Rechtfertigungen möglicherweise bereits an anderer Stelle formuliert wurden. Eine zeitlich orientierte Analyse könnte sichtbar machen, ob bestimmte Rechtfertigungen redundant werden oder sich etwa im Verlauf politischer Verhandlungen verhärten.

Die Analyse verweist außerdem auf konzeptionelle Limitationen und Forschungsbedarfe im Umgang mit der Cité der Solution. Die qualitative Herangehensweise ermöglichte es, den digitalen Geist eines optimistischen Determinismus sichtbar zu machen und so einen bislang wenig erfassten Teil dieser Rechtfertigungslogik zu beleuchten. Zugleich verliert die Cité der Solution aufgrund ihres diffusen Fortschrittsversprechens an Trennschärfe, da sie sich von anderen Cités durch eine mangelnde Verkörperung spezifischer Ansprüche unterscheidet. Stattdessen vertritt

sie einen optimistischen Determinismus, der verschiedene Nutzenversprechen gleichermaßen umfassen kann und damit beispielsweise auch Effizienz- und Wettbewerbsversprechen einschließen kann, sofern diese als gesellschaftliche Herausforderung betrachtet werden, die es durch den Einsatz von KI zu lösen gilt. Solutionistische Zukunftsvisionen erscheinen damit mitunter als eine Zuspitzung vergangener technologiepolitischer Effizienzversprechen oder auch als Querschnittslogik, die verschiedene moralische Ordnungen überlagern und verstärken kann. Eine solche Querschnittslogik hat das Potenzial, für verschiedenste Akteursgruppen anschlussfähig zu werden, wenn moralische Ansprüche unter dem Deckmantel des Solutionismus instrumentalisiert werden, ohne dass dies sichtbar und öffentlich verhandelt wird. Somit gilt es, zu untersuchen, ob solutionistische Narrative in verschiedenen Kontexten oder bei verschiedenen Akteur:innen unterschiedlich stark verfangen und ob eine mögliche Instrumentalisierung moralischer Ansprüche erkannt wird. Außerdem gilt es, die Verbreitung solutionistischer Deutungsmuster langfristig zu untersuchen, um mögliche Konflikte zu antizipieren sowie die Konsequenzen enttäuschter solutionistischer Erwartungen zu untersuchen.

Darauf aufbauend bleibt aufgrund des dokumentenanalytischen Zugangs ebenso offen, wie die in politischen Dokumenten formulierten Rechtfertigungen in der Öffentlichkeit wahrgenommen werden. In Anknüpfung an Suchmans (1995) Verständnis von Legitimität als Wahrnehmung der Angemessenheit politischer Entscheidungen stellt sich die Frage, ob die identifizierten Rechtfertigungen als überzeugend gelten oder ob sie etwa als Ausdruck symbolischer Politik wahrgenommen werden. Empirische Studien zur öffentlichen Wahrnehmung der vorgebrachten Rechtfertigungen können zeigen, wie verschiedene gesellschaftliche Gruppen die Legitimation des KI-Einsatzes im Bildungswesen bewerten und welche Erwartungen an eine demokratisch verantwortete KI-Governance bestehen. Damit ließe sich einschätzen, wie stabil die gegenwärtigen Legitimationspraktiken sind und inwieweit sie künftigen technologischen und politischen Herausforderungen standhalten.

11.3. Ausblick

Abschließend lässt sich die vorliegende Arbeit in einen übergeordneten gesellschaftlichen und technologiepolitischen Rahmen einordnen, der durch eine hohe Dynamik und Unsicherheit geprägt ist. Die Analyse der bildungspolitischen KI-Debatte zeigt, dass die Legitimation des KI-Einsatzes weit über Fragen technischer Funktionalität hinausgeht. Sie berührt grundlegende Vorstellungen darüber, wie Gesellschaft ihre Zukunft gestaltet, welche Werte sie priorisiert und welche Rolle technologische Instrumente in diesem Prozess einnehmen sollen. KI-Politik erweist sich damit als Spiegel gesellschaftlicher Aushandlungen über Gerechtigkeit, Verantwortung und das Verhältnis von Mensch und Technik.

Vor dem Hintergrund dieser Einordnung lässt sich eine vorsichtige Prognose wagen. In Anbetracht der politischen Steuerung wird die gesellschaftliche und politische Relevanz von KI weiter zunehmen, ebenso wie die Zahl der Anwendungsfelder. Mit der Ausweitung des KI-Einsatzes werden auch die moralischen Konflikte sichtbarer. Gallego und Kurer (2022) prognostizieren bereits eine zunehmende Politisierung von KI, insbesondere vor dem Hintergrund ungleich verteilter gesellschaftlicher Kosten und Nutzen. Diese Entwicklung dürfte sich im Bildungswesen ebenso fortsetzen wie in anderen sensiblen Bereichen. Bleibt die moralische Rechtfertigung politischer Entscheidungen jedoch diffus, technokratisch verkürzt oder auf symbolische Kommunikation reduziert, droht spätestens mit wirksam werdenden Risiken ein doppelter Legitimitätsverlust. Er betrifft sowohl die Technologie selbst als auch jene politischen Entscheidungsträger:innen, die ihren Einsatz verantworten. Dies kann gesellschaftliche Spaltungen vertiefen, besonders zwischen technologischen und politischen Eliten und jenen Gruppen, die von KI-Anwendungen betroffen sind, ohne die politischen Aushandlungen mitgestalten zu können. In diesem Zusammenhang kann die grundlegende Frage vertieft werden, ob KI in sensiblen Bereichen überhaupt eingesetzt werden sollte und welche Vorstellung einer guten und gerechten Gesellschaft politisches Handeln künftig leiten soll. Daran schließt sich die Debatte an, ob die Effizienzsteigerung tatsächlich einen erstrebenswerten Maßstab gesellschaftlicher Entwicklung darstellt oder ob andere moralische Orientierungen wie Gerechtigkeit, Fürsorge oder pädagogische Qualität stärker in den Vordergrund rücken sollten.

Die Analyse dieser Arbeit verweist darauf, dass Legitimität durch plurale moralische Rechtfertigungsordnungen erzeugt wird, deren Spannungen offen bearbeitet werden müssen. Technologische Entscheidungen können daher nicht allein auf Effizienzgewinne, Wettbewerbsfähigkeit oder Innovationsversprechen gestützt werden. Sie müssen eingebettet sein in eine politische Praxis, die Unsicherheit anerkennt, Konflikte sichtbar macht und Kritik ernst nimmt. Nur durch eine solche reflexive Auseinandersetzung kann vermieden werden, dass die KI-Entwicklung einem sich selbst verstärkenden Prozess folgt, der das technologische Momentum zu verpassen droht und in dem bestehende Pfadabhängigkeiten politische Alternativen unsichtbar machen, die eine Einführung von KI als vorbestimmt erscheinen lassen. Für die Zukunft der KI-Governance bedeutet dies, dass politische Entscheidungsprozesse inklusiver, geduldiger und stärker evidenzinformiert gestaltet werden müssen. Eine Governance, die kritische Perspektiven einbindet, moralische Ansprüche verschiedener gesellschaftlicher Gruppen berücksichtigt und ergebnisoffene Aushandlungsprozesse ermöglicht, kann dazu beitragen, die Legitimität des KI-Einsatzes zu sichern und zugleich gesellschaftliche Spannungen zu entlasten.

Die Arbeit zeigt, dass KI-Debatten einen zentralen Ort demokratischer Selbstverständigung darstellen. Sie eröffnen die Möglichkeit, grundlegende Fragen über gesellschaftliche Prioritäten, Gerechtigkeit und Verantwortung zu verhandeln und neu zu bestimmen. Die Zukunft des KI-Einsatzes im Bildungswesen und darüber hinaus hängt daher entscheidend davon ab, ob es gelingt, diesen Aushandlungsraum offen, kritisch und plural zu gestalten. In diesem Sinne bleibt die Gestaltung von KI nicht lediglich eine technische oder administrative Aufgabe, sondern eine fortlaufende gesellschaftliche Herausforderung, die demokratische Teilhabe, moralische Reflexion und politische Verantwortlichkeit erfordert.

Literaturverzeichnis

Abraham, U. (2025). Wie digital wollen wir werden? Das Verhältnis der Sozial- und Erziehungswissenschaften sowie der Deutschdidaktik zur Digitalisierung. In S. Emmersberger & I. Kammerer (Hrsg.), *Deutschdidaktik. Mediale Praktiken in einer Kultur der Digitalität* (S. 31–45). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-71769-1_2

Albrecht, J. P. (2016). How the GDPR Will Change the World. *European Data Protection Law Review*, *2*(3), 287–289.

Allhutter, D., Cech, F., Fischer, F., Grill, G. & Mager, A. (2020). Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective. *Frontiers in Big Data*, *3*(5), 1–17. https://doi.org/10.3389/fdata.2020.00005

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S. & Albekairy, A. M. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC medical education*, *23*(1), 1–15. https://doi.org/10.1186/s12909-023-04698-z

Altman, M., Wood, A. & Vayena, E. (2018). A Harm-Reduction Framework for Algorithmic Fairness. *IEEE Security & Privacy*, *16*(3), 34–45. https://doi.org/10.1109/MSP.2018.2701149

Alturki, S., Hulpuș, I. & Stuckenschmidt, H. (2022). Predicting Academic Outcomes: A Survey from 2007 Till 2018. *Technology, Knowledge and Learning*, *27*(1), 275–307. https://doi.org/10.1007/s10758-020-09476-0

Anesa, M., Spee, A. P., Gillespie, N. & Petani, F. J. (2022). Reassessing Moral Legitimacy in Times of Instability. *Journal of Management Studies*, 1–31. https://doi.org/10.1111/joms.12889

Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I. & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. *WIREs Data Mining and Knowledge Discovery*, *11*(5), Artikel e1424, 1–13. https://doi.org/10.1002/widm.1424

Angwin, J., Larson, J., Mattu, S. & Kirchner, L. (2016). *Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks.* ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Anicker, F. (2023). Gestaltwandel der Technik? Techniksoziologische Reflexionen zu künstlicher Intelligenz. In S. Alpsancar, A. Friedrich, P. Gehring, C. Hubig, A. Kaminski & A. Nordmann (Hrsg.), *Faktor Mensch: Jahrbuch Technikphilosophie 2023* (S. 137–168). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748941767-137

Anicker, F., Flaßhoff, G. & Marcinkowski, F. (2024). The Matrix of AI Agency: On the Demarcation Problem in Social Theory. *Sociological Theory*, *42*(4), 307–328. https://doi.org/10.1177/07352751241289925

Apetorgbor, M., Narad, S., Akpabio, E., Dhawale, C., Konto, A. & Verma, P. (2024). Leveraging Artificial Intelligence for Effective Assessment and Evaluation in Education: A Comprehensive Review. In *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI)* (S. 1–6). IEEE. https://doi.org/10.1109/IDICAIEI61867.2024.10842940

Attaran, M., Stark, J. & Stotler, D. (2018). Opportunities and challenges for big data analytics in US higher education. *Industry and Higher Education*, *32*(3), 169–182. https://doi.org/10.1177/0950422218770937

Awad, E., Dsouza, S., Bonnefon, J.-F., Shariff, A. & Rahwan, I. (2020). Crowdsourcing Moral Machines. *Communications of the ACM*, *63*(3), 48–55. https://doi.org/10.1145/3339904

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F. & Rahwan, I. (2018). The Moral Machine Experiment. *Nature*, *563*(7729), 59–64. https://doi.org/10.1038/s41586-018-0637-6

Baabdullah, A. M. (2024). The precursors of AI adoption in business: Towards an efficient decision-making and functional performance. *International Journal of Information Management*, *75*(1), Artikel 102745, 1–26. https://doi.org/10.1016/j.ijinfomgt.2023.102745

Bächtiger, A. & Wyss, D. (2013). Empirische Deliberationsforschung – eine systematische Übersicht. *Zeitschrift für Vergleichende Politikwissenschaft*, *7*(2), 155–181. https://doi.org/10.1007/s12286-013-0153-x

Baker, R. S. & Hawn, A. (2022). Algorithmic Bias in Education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

Ball, S. J. & Grimaldi, E. (2022). Neoliberal education and the neoliberal digital classroom. *Learning, Media and Technology*, *47*(2), 288–302. https://doi.org/10.1080/17439884.2021.1963980

Ballatore, A. & Natale, S. (2023). Technological failures, controversies and the myth of AI. In S. Lindgren (Hrsg.), *Handbook of critical studies of artificial intelligence* (S. 237–244). Edward Elgar Publishing.

Bao, L., Krause, N. M., Calice, M. N., Scheufele, D. A., Wirz, C. D., Brossard, D., Newman, T. P. & Xenos, M. A. (2022). Whose AI? How different publics think about AI and its social impacts. *Computers in Human Behavior*, *130*, 1–10. https://doi.org/10.1016/j.chb.2022.107182

Bareis, J. & Katzenbach, C. (2022). Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. *Science, Technology, & Human Values*, *47*(5), 855–881. https://doi.org/10.1177/01622439211030007

Barocas, S. & Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, *104*(3), 671–732.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R. & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Barry, A. (2001). *Political Machines: Governing a Technological Society*. Athlone Press.

Barry, M. & Blühdorn, I. (2018). Symbolic Environmental Politics. In N. Castree, M. Hulme & J. D. Proctor (Hrsg.), *Companion to Environmental Studies* (S. 249–253). Routledge.

Bartl, G. (2023). Krise und technologischer Solutionismus: Die politische Dimension des digitalisierten Umgangs mit Unsicherheit. In A. Wagener & C. Stark (Hrsg.), *Die Digitalisierung des Politischen: Theoretische und praktische Herausforderungen für die Demokratie* (S. 45–62). Springer VS. https://doi.org/10.1007/978-3-658-38268-1_3

Bartneck, C., Lütge, C., Wagner, A. & Welsh, S. (2021). *An Introduction to Ethics in Robotics and AI*. Springer International Publishing. https://doi.org/10.1007/978-3-030-51110-4

Bates, S., Jenkins, L. & Amery, F. (2014). (De)Politicisation and the Father's Clause Parliamentary Debates. *Policy & Politics*, *42*(2), 243–258. https://doi.org/10.1332/030557312X656007

Bauer, M. (1995). Resistance to New Technology and Its Effects on Nuclear Power, Information Technology and Biotechnology. In M. Bauer (Hrsg.), *Resistance to New Technology: Nuclear Power, Information Technology and Biotechnology* (S. 1–42). Cambridge University Press. https://doi.org/10.1017/CBO9780511563706.002

Beck, U. (2014). Wissen oder Nicht-Wissen? Zwei Perspektiven »reflexiver Modernisierung«. In U. Beck, A. Giddens & S. Lash (Hrsg.), *Reflexive Modernisierung: Eine Kontroverse* (S. 289–315). Suhrkamp Verlag.

Beck, U. (2016). *Risikogesellschaft: Auf dem Weg in eine andere Moderne* (23. Auflage). Suhrkamp. (Erstveröffentlichung 1986)

Beckman, L., Hultin Rosenberg, J. & Jebari, K. (2024). Artificial intelligence and democratic legitimacy. The problem of publicity in public authority. *AI & SOCIETY*, *39*(3), 975–984. https://doi.org/10.1007/s00146-022-01493-0

Beetham, D. (2013). *The Legitimation of Power* (2. Aufl.). Palgrave Macmillan. (Erstveröffentlichung 1991)

Bekker, S. (2021). Fundamental Rights in Digital Welfare States: The Case of SyRI in the Netherlands. In O. Spijkers, W. Werner & R. A. Wessel (Hrsg.), *Netherlands Yearbook of International Law 2019: Yearbooks in International Law: History, Function and Future* (Bd. 50, S. 289–307). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-403-7_24

Benkler, Y. (2019). Don't let industry write the rules for AI. *Nature*, *569*(7755), 161. https://doi.org/10.1038/d41586-019-01413-1

Bertsou, E. & Caramani, D. (2022). People Haven't Had Enough of Experts: Technocratic Attitudes among Citizens in Nine European Democracies. *American Journal of Political Science*, *66*(1), 5–23. https://doi.org/10.1111/ajps.12554

Bhatnagar, S., Alexandrova, A., Avin, S., Cave, S., Cheke, L., Crosby, M., Feyereisl, J., Halina, M., Loe, B. S., Ó hÉigeartaigh, S., Martínez-Plumed, F., Price, H., Shevlin, H., Weller, A., Winfield, A. & Hernández-Orallo, J. (2018). Mapping Intelligence: Requirements and Possibilities. In V. C. Müller (Hrsg.), *Philosophy and Theory of Artificial Intelligence 2017* (Bd. 44, S. 117–135). Springer International Publishing. https://doi.org/10.1007/978-3-319-96448-5_13

Bickerton, C. & Invernizzi Accetti, C. (2020). Technocracy and Political Theory. In E. Bertsou & D. Caramani (Hrsg.), *The Technocratic Challenge to Democracy* (S. 29–34). Routledge Taylor & Francis Group.

Bieber, D. & Deiß, M. (1997). Technikdeterminismus versus Gestaltungsautonomie: zum Einfluß von Technikmärkten auf die industrielle Rationalisierung. In D. Bieber, M. Deiß, H. Hirsch-Kreinsen & K. Schmierl (Hrsg.), *Neue Strukturen des Technikmarktes: Zur Entwicklung und Auslegung von Rechnersystemen für die industrielle Produktion. Institut für Sozialwissenschaftliche Forschung: ISF München* (S. 8–31).

Bijker, W. E. (1995). *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. MIT Press.

Bijker, W. E., Bal, R. & Hendriks, R. (2009). *The Paradox of Scientific Authority: The Role of Scientific Advice in Democracies*. *Inside technology*. The MIT Press.

Blatter, J. (2007). Demokratie und Legitimation. In A. Benz, S. Lütz, U. Schimank & G. Simonis (Hrsg.), *Handbuch Governance: Theoretische Grundlagen und empirische Anwendungsfelder* (S. 271–284). VS Verlag für Sozialwissenschaften.

Blättler, S. (2008). Stichwort: Luc Boltanski - Eine Soziologie der Moral. *WestEnd - Neue Zeitschrift für Sozialforschung*, *5*(2), 79–83.

Bleher, H. & Braun, M. (2023). Reflections on Putting AI Ethics into Practice: How Three AI Ethics Approaches Conceptualize Theory and Practice. *Science and engineering ethics*, *29*(3), 21. https://doi.org/10.1007/s11948-023-00443-3

Bleiklie, I. (2018). New Public Management or Neoliberalism, Higher Education. In P. N. Teixeira & J.-C. Shin (Hrsg.), *Encyclopedia of International Higher Education Systems and Institutions* (S. 1–6). Springer Netherlands.

Blühdorn, I. (2007). Sustaining the unsustainable: Symbolic politics and the politics of simulation. *Environmental Politics*, *16*(2), 251–275. https://doi.org/10.1080/09644010701211759

Blümel, C. (2021). Innovations in innovation policy: reconstructing the emergence, legitimation and dynamics of cluster policies in Germany. *Innovation: The European Journal of Social Science Research*, *34*(4), 533–559. https://doi.org/10.1080/13511610.2020.1798216

BMBF. (2016). *Richtlinie zur Förderung von Forschung zur digitalen Hochschulbildung – Wirksamkeit und Wirkungen aktueller Ansätze und Formate – Trends und neue Paradigmen in Didaktik und Technik. Bundesanzeiger vom 26.02.2016.* Bundesministerium für Bildung und Forschung. https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2016/02/1152_bekanntmachung

Boella, G., Di Caro, L., Humphreys, L., Robaldo, L., Rossi, P. & van der Torre, L. (2016). Eunomos, a legal document and knowledge management system for the Web to provide relevant, reliable and up-to-date information on the law. *Artificial Intelligence and Law*, *24*(3), 245–283. https://doi.org/10.1007/s10506-016-9184-3

Bogner, A. (2021a). *Die Epistemisierung des Politischen: Wie die Macht des Wissens die Demokratie gefährdet*. Reclam.

Bogner, A. (2021b). Politisierung, Demokratisierung, Pragmatisierung. Paradigmen der Technikfolgenabschätzung im Wandel der Zeit. In S. Böschen, A. Grunwald, B.-J. Krings & C. Rösch (Hrsg.), *Technikfolgenabschätzung: Handbuch für Wissenschaft und Praxis* (S. 43–58). Nomos Verlagsgesellschaft mbH & Co. KG.

Bogusz, T. (2010). *Zur Aktualität von Luc Boltanski: Einleitung in sein Werk*. VS Verlag für Sozialwissenschaften.

Bogusz, T. (2013). Was heißt Pragmatismus? Boltanski meets Dewey. *Berliner Journal für Soziologie*, *23*(3-4), 311–328. https://doi.org/10.1007/s11609-013-0226-3

Bohnsack, R., Nentwig-Gesemann, I. & Nohl, A.-M. (Hrsg.). (2013a). *Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-19895-8

Bohnsack, R., Nentwig-Gesemann, I. & Nohl, A.-M. (2013b). Einleitung: Die dokumentarische Methode und ihre Forschungspraxis. In R. Bohnsack, I. Nentwig-Gesemann & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung* (S. 9–32). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-19895-8_1

Böker, A. (2021). *Über die Rechtfertigung von Begabtenförderung*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-31529-0

Bolsen, T. & Druckman, J. N. (2015). Counteracting the Politicization of Science. *Journal of Communication*, *65*(5), 745–769. https://doi.org/10.1111/jcom.12171

Boltanski, L. (2008). Individualismus ohne Freiheit. Ein pragmatischer Zugang zur Herrschaft. *WestEnd - Neue Zeitschrift für Sozialforschung*, *5*(2), 133–149.

Boltanski, L. (2010). *Soziologie und Sozialkritik. Frankfurter Adorno-Vorlesungen*. Suhrkamp.

Boltanski, L. (2011). *On Critique: A Sociology of Emancipation*. Polity Press.

Boltanski, L. & Chiapello, È. (1999). *Le Nouvel Esprit du Capitalisme*. Gallimard.

Boltanski, L. & Chiapello, E. (2001). Die Rolle der Kritik in der Dynamik des Kapitalismus und der normative Wandel. *Berliner Journal für Soziologie*, *11*(4), 459–477. https://doi.org/10.1007/BF03204032

Boltanski, L. & Chiapello, È. (2003). *Der neue Geist des Kapitalismus. Édition Discours: Bd. 30*. UVK-Verlagsgesellschaft.

Boltanski, L. & Thévenot, L. (1991). *De la Justification: Les Économies de la Grandeur. NRF essais*. Gallimard.

Boltanski, L. & Thévenot, L. (2018). *Über die Rechtfertigung: Eine Soziologie der kritischen Urteilskraft* (2. Auflage). Hamburger Edition. (Erstveröffentlichung 1991)

Booth, A., Sutton, A. & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review* (Second edition). SAGE.

Borup, M., Brown, N., Konrad, K. & van Lente, H. (2006). The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management*, *18*(3-4), 285–298. https://doi.org/10.1080/09537320600777002

Bouza García, L. & Oleart, A. (2024). Regulating Disinformation and Big Tech in the EU: A Research Agenda on the Institutional Strategies, Public Spheres and Analytical Challenges. *JCMS: Journal of Common Market Studies, 62*(5), 1395–1407. https://doi.org/10.1111/jcms.13548

boyd, d. & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Boyd, M. & Wilson, N. (2017). Rapid Developments in Artificial Intelligence: How Might the New Zealand Government Respond? *Policy Quarterly*, *13*(4), 36–43. https://doi.org/10.26686/pq.v13i4.4619

Boyd, R. & Holton, R. J. (2018). Technology, innovation, employment and power: Does robotics and artificial intelligence really mean social transformation? *Journal of Sociology*, *54*(3), 331–345. https://doi.org/10.1177/1440783317726591

Brantner, C. & Saurwein, F. (2021). Covering Technology Risks and Responsibility: Automation, Artificial Intelligence, Robotics, and Algorithms in the Media. *International Journal of Communication*, *15*(1), 5074–5098.

Brauner, P., Glawe, F., Liehner, G. L., Vervier, L. & Ziefle, M. (2025). Mapping public perception of artificial intelligence: Expectations, risk–benefit tradeoffs, and value as determinants for societal acceptance. *Technological Forecasting and Social Change*, *220*, 1–17. https://doi.org/10.1016/j.techfore.2025.124304

Brause, S. R., Schäfer, M. S., Katzenbach, C., Mao, Y., Richter, V. & Zeng, J. (2025). Sociotechnical imaginaries and public communication: Analytical framework and empirical illustration using the case of artificial intelligence. Convergence: The International *Journal of Research into New Media Technologies, 31*(4), 1267–1287. https://doi.org/10.1177/13548565251338192

Brennen, J. S., Howard, P. N. & Nielsen, R. K [Rasmus Kleis]. (2018). *An Industry-Led Debate: How UK Media Cover Artificial Intelligence*. Factsheet. Reuters Institute.

Brennen, J. S., Howard, P. N. & Nielsen, R. K [Rasmus K.] (2022). What to Expect When You're Expecting Robots: Futures, Expectations, and Pseudo-Artificial General Intelligence in UK News. *Journalism*, *23*(1), 22–38. https://doi.org/10.1177/1464884920947535

Brockmann, H., Drews, W. & Torpey, J. (2021). A Class for Itself? On the Worldviews of the New Tech Elite. *PloS ONE*, *16*(1), 1-26. https://doi.org/10.1371/journal.pone.0244071

Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*. *The MIT Press*. The MIT Press.

Brown, M. B. (2015). Politicizing science: conceptions of politics in science and technology studies. *Social Studies of Science*, *45*(1), 3–30. https://doi.org/10.1177/0306312714556694

Brown, N. (2003). Hope Against Hype - Accountability in Biopasts, Presents and Futures. *Science & Technology Studies*, *16*(2), 3–21. https://doi.org/10.23987/sts.55152

Brynjolfsson, E. & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company Incorporated.

Brynjolfsson, E., Rock, D. & Syverson, C. (2017). *Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics*. NBER Working Paper No. 24001. https://doi.org/10.3386/w24001

Büching, C., Mah, D.-K., Otto, S., Paulicke, P. & Hartman, E. A. (2019). Learning Analytics an Hochschulen. In V. Wittpahl (Hrsg.), *Künstliche Intelligenz: Technologie - Anwendung - Gesellschaft* (S. 142–160). Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-58042-4_9

Bundesregierung. (2018). *Strategie Künstliche Intelligenz der Bundesregierung*. https://www.ki-strategie-deutschland.de/files/downloads/Nationale_KI-Strategie.pdf

Bundesregierung. (2020). *Strategie Künstliche Intelligenz der Bundesregierung: Fortschreibung 2020*. https://www.ki-strategie-deutschland.de/files/downloads/201201_Fortschreibung_KI-Strategie.pdf

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), Artikel 1-12. https://doi.org/10.1177/2053951715622512

Burrell, J. & Fourcade, M. (2021). The Society of Algorithms. *Annual Review of Sociology*, *47*(1), 213–237. https://doi.org/10.1146/annurev-soc-090820-020800

Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. *Public Administration Review*, *81*(5), 825–836. https://doi.org/10.1111/puar.13293

Butcher, J. & Beridze, I. (2019). What is the State of Artificial Intelligence Governance Globally? *The RUSI Journal*, *164*(5-6), 88–96. https://doi.org/10.1080/03071847.2019.1694260

Callaghan, C. W. (2018). Surviving a Technological Future: Technological Proliferation and Modes of Discovery. *Futures*, *104*, 100–116. https://doi.org/10.1016/j.futures.2018.08.001

Callon, M. & Law, J. (1995). Agency and the Hybrid Collectif. *The South Atlantic Quarterly*, *94*(2), 481–507.

Campolo, A. & Crawford, K. (2020). Enchanted Determinism: Power without Responsibility in Artificial Intelligence. *Engaging Science, Technology, and Society*, *6*, 1–19. https://doi.org/10.17351/ests2020.277

Cantero Gamito, M. & Ebers, M. (2021). Algorithmic Governance and Governance of Algorithms: An Introduction. In M. Ebers & M. Cantero Gamito (Hrsg.), *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges* (Bd. 1, S. 1–22). Springer International Publishing. https://doi.org/10.1007/978-3-030-50559-2_1

Cao, L. (2023). AI in Finance: Challenges, Techniques, and Opportunities. *ACM Computing Surveys*, *55*(3), 1–38. https://doi.org/10.1145/3502289

Capano, G. & Woo, J. J. (2017). Resilience and robustness in policy design: a critical appraisal. *Policy Sciences*, *50*(3), 399–426. https://doi.org/10.1007/s11077-016-9273-x

Caramani, D. (2020). Introduction: the Technocratic Challenge to Democracy. In E. Bertsou & D. Caramani (Hrsg.), *The Technocratic Challenge to Democracy* (S. 1–26). Routledge Taylor & Francis Group.

Cardno, C. (2018). Policy Document Analysis: A Practical Educational Leadership Tool and a Qualitative Research Method. *Educational Administration: Theory and Practice*, *24*(4), 623–640. https://doi.org/10.14527/kuey.2018.016

Carpentier, V. & Courtois, A. (2025). Higher education and the public good in France. *Higher Education*, *89*(1), 223–238. https://doi.org/10.1007/s10734-024-01293-z

Castañeda, L. & Selwyn, N. (2018). More than tools? Making sense of the ongoing digitizations of higher education. *International Journal of Educational Technology in Higher Education*, *15*(1), 1–10. https://doi.org/10.1186/s41239-018-0109-y

Cath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133), 1–8. https://doi.org/10.1098/rsta.2018.0080

Cave, S., Coughlan, K. & Dihal, K. (2019). "Scary Robots": Examining Public Responses to AI. In V. Conitzer, G. Hadfield & S. Vallor (Hrsg.), *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (S. 331–337). ACM. https://doi.org/10.1145/3306618.3314232

Cave, S. & Dihal, K. (2019). Hopes and fears for intelligent machines in fiction and reality. *Nature Machine Intelligence*, *1*(2), 74–78. https://doi.org/10.1038/s42256-019-0020-9

Chai, S., Doshi, A. R. & Silvestri, L. (2022). How Catastrophic Innovation Failure Affects Organizational and Industry Legitimacy: The 2014 Virgin Galactic Test Flight Crash. *Organization Science, 33*(3), 1068–1093. https://doi.org/10.1287/orsc.2021.1467

Chai, F., Ma, J., Wang, Y., Zhu, J. & Han, T. (2024). Grading by AI makes me feel fairer? How different evaluators affect college students' perception of fairness. *Frontiers in psychology*, *15*, 1–18. https://doi.org/10.3389/fpsyg.2024.1221177

Cheatham, B., Javanmardian, K. & Samandari, H. (2019). Confronting the risks of artificial intelligence: With great power comes great responsibility. Organizations can mitigate the risks of applying artificial intelligence and advanced analytics by embracing three principles. *McKinsey Quarterly*(2), 1–9.

Chen, B., Zhu, X. & Del Díaz Castillo H., F. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence, 5*, 1–12. https://doi.org/10.1016/j.caeai.2023.100184

Chesney, B. & Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, *107*(1), 1753–1819. https://doi.org/10.15779/Z38RVOD15J

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S. & Cheng, M. (2023). Systematic Literature Review on Opportunities, Challenges, and Future Research Recommendations of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *4*, 1–15. https://doi.org/10.1016/j.caeai.2022.100118

Christen, M., Mader, C., Čas, J., Abou-Chadi, T., Bernstein, A., Braun Binder, N., Dell'Aglio, D., Fábián, L., George, D., Gohdes, A., Hilty, L. M., Kneer, M., Krieger-Lamina, J., Licht, H., Scherer, A., Som, C., Sutter, P. & Thouvenin, F. (2020). *Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz*. *TA-Swiss: 72 (2020)*. vdf. https://doi.org/10.3218/4002-9

Chuan, C.-H., Tsai, W.-H. S. & Cho, S. Y. (2019). Framing Artificial Intelligence in American Newspapers. In V. Conitzer, G. Hadfield & S. Vallor (Hrsg.), *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (S. 339–344). ACM. https://doi.org/10.1145/3306618.3314285

Clayton, E. W. (2003). Ethical, Legal, and Social Implications of Genomic Medicine. *The New England journal of medicine*, *349*(6), 562–569. https://doi.org/10.1056/NEJMra012577

Cohen, J [Jacob] (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J [Joshua]. (2003). Deliberation and Democratic Legitimacy. In D. Matravers & J. Pike (Hrsg.), *Debates in Contemporary Political Philosophy: An Anthology* (S. 342–360). Routledge.

Cohen, J. E. (2019). *Between Truth and Power: The Legal Constructions of Informational Capitalism*. Oxford University Press.

Collingridge, D. (1985). Controlling Technology (Response to Johnston). *Social Studies of Science*, *15*(2), 373–380.

Conti, L. G. & Seele, P. (2023). The contested role of AI ethics boards in smart societies: a step towards improvement based on board composition by sortition. *Ethics and Information Technology*, *25*(4), 1–15. https://doi.org/10.1007/s10676-023-09724-8

Cools, H., van Gorp, B. & Opgenhaffen, M. (2024). Where exactly between utopia and dystopia? A framing analysis of AI and automation in US newspapers. *Journalism*, *25*(1), 3–21. https://doi.org/10.1177/14648849221122647

Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. (2022). *Introduction to Algorithms* (4. Aufl.). The MIT Press. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6925615

Corrêa, N. K., Galvão, C., Santos, J. W., Del Pino, C., Pinto, E. P., Barbosa, C., Massmann, D., Mambrini, R., Galvão, L., Terem, E. & Oliveira, N. de (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, *4*(10), 1–14. https://doi.org/10.1016/j.patter.2023.100857

Couldry, N. (2020). Recovering critique in an age of datafication. *New Media & Society*, *22*(7), 1135–1151. https://doi.org/10.1177/1461444820912536

Cowls, J., Tsamados, A., Taddeo, M. & Floridi, L. (2023). The AI gambit: leveraging artificial intelligence to combat climate change-opportunities, challenges, and recommendations. *AI & SOCIETY*, *38*(1), 283–307. https://doi.org/10.1007/s00146-021-01294-x

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.* Yale University Press.

Creely, E. & Blannin, J. (2025). Creative partnerships with generative AI. Possibilities for education and beyond. *Thinking Skills and Creativity, 56*, 1–15. https://doi.org/10.1016/j.tsc.2024.101727

Creswell, J. W. & Creswell Báez, J. (2021). *30 essential skills for the qualitative researcher* (Second edition). SAGE.

Csernatoni, R. (2020). New States of Emergency: Normalizing Techno-Surveillance in the Time of COVID-19. *Global Affairs*, *6*(3), 301–310. https://doi.org/10.1080/23340460.2020.1825108

Dafoe, A. (2015). On Technological Determinism: A Typology, Scope Conditions, and a Mechanism. *Science, Technology, & Human Values*, *40*(6), 1047–1076. https://doi.org/10.1177/0162243915579283

Danaher, J. (2016). The Threat of Algocracy: Reality, Resistance and Accommodation. *Philosophy & Technology*, *29*(3), 245–268. https://doi.org/10.1007/s13347-015-0211-1

Danaher, J., Hogan, M. J., Noone, C., Kennedy, R., Behan, A., Paor, A. de, Felzmann, H., Haklay, M., Khoo, S.-M., Morison, J., Murphy, M. H., O'Brolchain, N., Schafer, B. & Shankar, K. (2017). Algorithmic Governance: Developing a Research Agenda Through the Power of Collective Intelligence. *Big Data & Society*, *4*(2), 1-21. https://doi.org/10.1177/2053951717726554

Dastin, J. (2022). Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women. In K. Martin (Hrsg.), *Ethics of Data and Analytics: Concepts and Cases* (S. 296–299). Auerbach Publications. https://doi.org/10.1201/9781003278290-44

Datenethikkommission der Bundesregierung (2019). *Gutachten der Datenethikkommission der Bundesregierung.* Datenenthikkommission der Bundesregierung; Bundesministerium des Innern, für Bau und Heimat; Bundesministerium der Justiz und für Verbraucherschutz. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=8

Davidow, B. (20. Februar 2014). Welcome to Algorithmic Prison: The use of Big Data to to profile citizens is subtly, silently constraining freedom. *The Atlantic*. https://www.theatlantic.com/technology/archive/2014/02/welcome-to-algorithmic-prison/283985/

de Bruijn, H., Warnier, M. & Janssen, M. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly*, *39*(2), 1–8. https://doi.org/10.1016/j.giq.2021.101666

de Fine Licht, J. (2020). The Janus Face of Transparency: Balancing Openness and Secrecy in Democrativ Decision-Making. In D. Mokrosinska (Hrsg.), *Transparency and Secrecy in European Democracies: Contested Trade-offs* (S. 17–35). Routledge.

de Fine Licht, K. & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making: Why explanations are key when trying to produce perceived legitimacy. *AI & SOCIETY*, *35*(4), 917–926. https://doi.org/10.1007/s00146-020-00960-w

de Laat, P. B. (2018). Algorithmic Decision-Making Based on Machine Learning from Big Data: Can Transparency Restore Accountability? *Philosophy & Technology*, *31*(4), 525–541. https://doi.org/10.1007/s13347-017-0293-z

de Wilde, P. (2011). No Polity for Old Politics? A Framework for Analyzing the Politicization of European Integration. *Journal of European Integration*, *33*(5), 559–575. https://doi.org/10.1080/07036337.2010.546849

de Wilde, P., Leupold, A. & Schmidtke, H. (2016). Introduction: the differentiated politicisation of European governance. *West European Politics*, *39*(1), 3–22. https://doi.org/10.1080/01402382.2015.1081505

Degele, N. (2002). *Einführung in die Techniksoziologie*. UTB.

Denzin, N. K. (2017). *The Research Act: A Theoretical Introduction to Sociological Methods*. Routledge. https://doi.org/10.4324/9781315134543 (Erstveröffentlichung 1970)

Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, *132*(20), 1920–1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

Desrosières, A. (1998). *The Politics of Large Numbers: A History of Statistical Reasoning*. Harvard University Press.

Desrosières, A. (2001). How Real Are Statistics? Four Posssible Attitudes. *Social Research*, *68*(2), 339–355.

Desrosières, A. (2015). Retroaction: How Indicators Feed Back onto Quantified Actors. In R. Rottenburg, S. E. Merry, S.-J. Park & J. Mugler (Hrsg.), *The World of Indicators: The Making of Governmental Knowledge through Quantification* (S. 29–353). Cambridge University Press.

Desrosières, A. (2025). Die Statistik: Instrument der Befreiung oder Instrument der Macht? (S. Gruhl, Übers.). *À propos*(1), 1–22. https://doi.org/10.57086/apropos.183 (Erstveröffentlichung 2014)

Dessauer, F. (1927). *Philosophie der Technik: das Problem der Realisierung*. Cohen.

Diakopoulos, N. (2016). Accountability in Algorithmic Decision Making. *Communications of the ACM*, *59*(2), 56–62. https://doi.org/10.1145/2844110

Diaz-Bone, R. (2018). *Die „Economie des conventions“: Grundlagen und Entwicklungen der neuen französischen Wirtschaftssoziolgie* (2. Aufl.). Springer VS. https://doi.org/10.1007/978-3-658-21062-5

Diaz-Bone, R. (2022). Soziologie der Konventionen. In H. Delitz (Hrsg.), *Soziologische Denkweisen aus Frankreich* (S. 471–493). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36949-1_18

Diaz-Bone, R. & Didier, E. (2016). The Sociology of Quantification - Perspectives on an Emerging Field in the Social Sciences. *Historical Social Research*, *41*(2), 7–26. https://doi.org/10.12759/hsr.41.2016.2.7.7-26

Diaz-Bone, R. & Horvath, K. (2021). Official Statistics, Big Data and Civil Society. Introducing the Approach of "Economics of Convention” for Understanding the Rise of New Data Worlds and Their Implications. *Statistical Journal of the IAOS, 37*(1), 219–228. https://doi.org/10.3233/SJI-200733

Diaz-Bone, R., Horvath, K. & Cappel, V. (2020). Social Research in Times of Big Data. The Challenges of New Data Worlds and the Need for a Sociology of Social Research. *Historical Social Research, 45*(3), 314–342. https://doi.org/10.12759/HSR.45.2020.3.314-341

Diaz-Bone, R. & Schrör, S. (2022). Algorithms, Conventions and New Regulation Processes. In M. Filimowicz (Hrsg.), *Democratic Frontiers: Algorithms and Society* (S. 24–46). Routledge. https://doi.org/10.4324/9781003173427-2

Diehl, P. (2015). Zum Zusammenhang von Legitimität, Legitimation und symbolischer Repräsentation. In R. Voigt (Hrsg.), *Legalität ohne Legitimität? Carl Schmitts Kategorie der Legitimität* (S. 281–296). Springer VS.

Diekmann, A. (2007). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Rowohlt Taschenbuch Verlag.

Ding, I. (2020). Performative Governance. *World Politics*, *72*(4), 525–556. https://doi.org/10.1017/S0043887120000131

Doezema, T. & Frahm, N. (2023). The New Spirit of Technoscience: Recalibrating Symmetrical STS Critique. *Journal of Responsible Innovation*, *10*(1), 1–12. https://doi.org/10.1080/23299460.2023.2281112

Dogruel, L., Facciorusso, D. & Stark, B. (2022). ‘I’m still the master of the machine.’ Internet users’ awareness of algorithmic decision-making and their perception of its effect on their autonomy. *Information, Communication & Society*, *25*(9), 1311–1332. https://doi.org/10.1080/1369118X.2020.1863999

Dolata, U. (2008). Technologische Innovationen und sektoraler Wandel: Eingriffstiefe, Adaptionsfähigkeit, Transformationsmuster: Ein analytischer Ansatz. *Zeitschrift für Soziologie*, *37*(1), 42–59.

Dolata, U. & Werle, R. (2007). „Bringing Technology Back In“: Technik als Einflussfaktor sozioökonomischen und institutionellen Wandels. In U. Dolata & R. Werle (Hrsg.), *Gesellschaft und die Macht der Technik: Sozioökonomischer und Institutioneller Wandel durch Technisierung* (S. 15–44). Campus Verlag.

Du, J. (2024). The Impact of Artificial Intelligence Adoption on Employee Unemployment: A Multifaceted Relationship. *International Journal of Social Sciences and Public Administration*, *2*(3), 321–327. https://doi.org/10.62051/ijsspa.v2n3.45

Duan, Y., Edwards, J. S. & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, *48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Dumbach, P., Schwinn, L., Löhr, T., Elsberger, T. & Eskofier, B. M. (2023). Artificial intelligence trend analysis in German business and politics: a web mining approach. *International Journal of Data Science and Analytics*, 1–22. https://doi.org/10.1007/s41060-023-00483-9

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., . . . Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*(57), 1–47. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Easton, D. (1965). *A systems analysis of political life*. Wiley.

Easton, D. (1975). A Re-assessment of the Concept of Political Support. *British Journal of Political Science*, *5*(4), 435–457. https://doi.org/10.1017/S0007123400008309

Edelman, M. J. (2005). *Politik als Ritual: Die symbolische Funktion staatlicher Institutionen und politischen Handelns* (3., erw. Aufl.). Campus-Verlag. (Erstveröffentlichung 1976)

Eden, C. A., Chisom, O. N. & Adeniyi, I. S. (2024). Integrating AI in education: Opportunities, challenges, and ethical considerations. *Magna Scientia Advanced Research and Reviews*, *10*(2), 6–13. https://doi.org/10.30574/msarr.2024.10.2.0039

Edler, J. & Wittmann, F. (2024). Technology policy and governance. In A. Grunwald (Hrsg.), *Handbook of Technology Assessment* (S. 12–21). Edward Elgar Publishing.

Edwards, C. (2021). Let the algorithm decide? *Communications of the ACM*, *64*(6), 21–22. https://doi.org/10.1145/3460216

Elhaddad, M. & Hamam, S. (2024). AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential. *Cureus*, *16*(4), 1-9. https://doi.org/10.7759/cureus.57728

Elish, M. C. & boyd, d. (2018). Situating Methods in the Magic of Big Data and AI. *Communication Monographs*, *85*(1), 57–80. https://doi.org/10.1080/03637751.2017.1375130

Ellul, J. (2021). *The Technological Society*. Knopf Doubleday Publishing Group. (Erstveröffentlichung 1954)

Engin, Z. & Treleaven, P. (2019). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. *The Computer Journal*, *62*(3), 448–460. https://doi.org/10.1093/comjnl/bxy082

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication, 43*(4), 51–58.

Espeland, W. N. & Stevens, M. L. (2008). A Sociology of Quantification. *European Journal of Sociology*, *49*(3), 401–436. https://doi.org/10.1017/S0003975609000150

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin‘s Press.

Europäische Kommission (2019). *Ethik-Leitlinien für eine vertrauenswürdige KI.* Hochrangige Expertengruppe für künstliche Intelligenz; Europäische Kommission. https://doi.org/10.2759/22710

Europäische Union (2014). *Rome Declaration on Responsible Research and Innovation in Europe*. https://digital-strategy.ec.europa.eu/de/node/7302

Europäische Union (2016). Verordnung 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 - zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/ 46/ EG (Datenschutz-Grundverordnung).

Europäische Union (2024). Verordnung 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) (2024).

Everson, M. (2020). Politicising Europe: liberating the technocratic? In P. J. Cardwell & M.-P. Granger (Hrsg.), *Research handbooks in lLaw and politics. Research handbook on the politics of EU law* (S. 144–161). Edward Elgar Publishing. https://doi.org/10.4337/9781788971287.00017

Eynon, R. & Young, E. (2021). Methodology, Legend, and Rhetoric: The Constructions of AI by Academia, Industry, and Policy Groups for Lifelong Learning. *Science, Technology, & Human Values*, *46*(1), 166–191. https://doi.org/10.1177/0162243920906475

Fatt, S. J. & Fardouly, J. (2025). Debate: Social media in children and young people – time for a ban? Weighing up the implications and limitations of age-based social media restrictions. *Child and Adolescent Mental Health, 30*(4), 414–415. https://doi.org/10.1111/camh.70034

Fast, E. & Horvitz, E. (2017). Long-Term Trends in the Public Perception of Artificial Intelligence. In S. Singh, J. Crawford & E. Eaton (Hrsg.), *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)* (S. 963–969). AAAI Press. https://doi.org/10.48550/arXiv.1609.04904

Fawcett, P., Flinders, M., Hay, C. & Wood, M. (Hrsg.). (2017). *Anti-Politics, Depoliticization, and Governance*. Oxford University Press.

Fazelpour, S. & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, *16*(8), 1–16. https://doi.org/10.1111/phc3.12760

Feenberg, A. (1991). *Critical Theory of Technology*. Oxford University Press.

Feenberg, A. (1999). *Questioning Technology*. Routledge. https://doi.org/10.4324/9780203022313

Feenberg, A. (2002). *Transforming Technology: A Critical Theory Revisited* (Revised ed.). Oxford University Press. https://doi.org/21399

Feenberg, A. (2009). Critical Theory of Technology: An Overview. In G. J. Leckie & J. E. Buschman (Hrsg.), *Information Technology in Librarianship: New Critical Approaches* (S. 31–46). Libraries Unlimited.

Feldstein, S. (2019). How Artificial Intelligence is Reshaping Repression. *Journal of Democracy*, *30*(1), 40–52. https://doi.org/10.1353/jod.2019.0003

Felzmann, H., Villaronga, E. F., Lutz, C. & Tamò-Larrieux, A. (2019). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society*, *6*(1), Artikel 2053951719860542. https://doi.org/10.1177/2053951719860542

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B. & Ullmann, Thomas and Vuorikari, Riina. (2016). *Research Evidence on the Use of Learning Analytics: Implications for Education Policy.* Joint Research Centre Science for Policy Reports. https://doi.org/10.2791/955210

Ferrari, E. (2020). Technocracy Meets Populism: The Dominant Technological Imaginary of Silicon Valley. *Communication, Culture and Critique*, *13*(1), 121–124. https://doi.org/10.1093/ccc/tcz051

Fischer, S. & Puschmann, C. (2021). *Wie Deutschland über Algorithmen schreibt: Eine Analyse des Mediendiskurses über Algorithmen und Künstliche Intelligenz (2005-2020).* Bertelsmann Stiftung. https://doi.org/10.11586/2021003

Fisher, E., Mahajan, R. L. & Mitcham, C. (2006). Midstream Modulation of Technology: Governance From Within. *Bulletin of Science, Technology & Society*, *26*(6), 485–496. https://doi.org/10.1177/0270467606295402

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. C. & Srikumar, M. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI.* The Berkman Klein Center for Internet & Society Research Publication Series. https://doi.org/10.2139/ssrn.3518482

Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications*, *7*(1), 1–9. https://doi.org/10.1057/s41599-020-0494-4

Flick, U. (2020). Gütekriterien qualitativer Forschung. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 247–263). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-26887-9_30

Ford, M. (2015). *Rise of the Robots: Technology and the Threat of a Jobless Future*. Basic Books.

Forst, R. (2014). Legitimität, Demokratie und Gerechtigkeit: Zur Reflexivität normativer Ordnungen. In O. Flügel-Martinsen, D. Gaus, T. Hitzel-Cassagnes & F. Martinsen (Hrsg.), *Deliberative Kritik - Kritik der Deliberation: Festschrift für Rainer Schmalz-Bruns* (S. 137–147). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-02860-2_6

Forst, R. (2015). *Normativität und Macht: Zur Analyse sozialer Rechtfertigungsordnungen*. Suhrkamp.

Forst, R. (2018). The Ground of Critique: On the Concept of Human Dignity in Social Orders of Justification. In A. Nascimento & M. Lutz-Bachmann (Hrsg.), *Human Dignity: Perspectives from a Critical Theory of Human Rights* (S. 71–83). Routledge.

Forst, R. & Günther, K. (2011). Die Herausbildung normativer Ordnungen: Zur Idee eines interdisziplinären Forschungsprogramms. In R. Forst & K. Günther (Hrsg.), *Die Herausbildung normativer Ordnungen: Interdisziplinäre Perspektiven* (S. 11–30). Campus Verlag.

Foucault, M. (2024). *Überwachen und Strafen: Die Geburt des Gefängnisses* (W. Seitter, Übers.) (20. Auflage). Suhrkamp. (Erstveröffentlichung 1975)

Fourcade, M. & Healy, K. (2024). *The Ordinal Society*. Harvard University Press. https://doi.org/10.4159/9780674296688

Frahm, N., Doezema, T. & Pfotenhauer, S. (2022). Fixing Technology with Society: The Coproduction of Democratic Deficits and Responsible Innovation at the OECD and the European Commission. *Science, Technology, & Human Values*, *47*(1), 174–216. https://doi.org/10.1177/0162243921999100

Frahm, N. & Schiølin, K. H. (2023). Toward an 'Ever Closer Union': The Making of AI-Ethics in the EU. *STS Encounters*, *15*(2), 2–25. https://doi.org/10.7146/stse.v15i2.139808

Frei, A. I., Steinberg, M. & Horvath, K. (2023). Algorithmische Sortierungen: Eine neopragmatische Perspektive auf KI-gestützte Klassifikationen in der schulischen Bildung. In M. Schiefner-Rohs, S. Hofhues & A. Breiter (Hrsg.), *Datafizierung (in) der Bildung: Kritische Perspektiven auf digitale Vermessung in pädagogischen Kontexten* (S. 131–154). transcript. https://doi.org/10.14361/9783839465820-009

Freyer, H. (1987). Über das Dominantwerden technischer Kategorien in der Lebenswelt der industriellen Gesellschaft. In E. Üner (Hrsg.), *Herrschaft, Planung und Technik: Aufsätze zur politischen Soziologie* (S. 117–130). VCH Acta Humaniora.

Frouillou, L., Pin, C. & van Zanten, A. (2020). Have the APB and Parcoursup platforms promoted equal opportunity? The evolution of standards and procedures for access to higher education in France. *L'Année Sociologique*, *70*(2), 337–363. https://doi.org/10.3917/anso.202.0337

Future of Life Institute (2023). *Pause Giant AI Experiments: An Open Letter*. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Gadinger, F. (2016). On Justification and Critique: Luc Boltanski's Pragmatic Sociology and International Relations. *International Political Sociology*, *10*(3), 187–205. https://doi.org/10.1093/ips/olw007

Gadinger, F. & Yildiz, T. (2012). Unruhige Legitimität. Die Finanzkrise aus Sicht der pragmatischen Soziologie der Kritik. In A. Geis, F. Nullmeier & C. Daase (Hrsg.), *Der Aufstieg der Legitimitätspolitik: Rechtfertigung und Kritik politisch-ökonomischer Ordnungen* (S. 302–319). Nomos. https://doi.org/10.5771/9783845245072-302

Gallego, A. & Kurer, T. (2022). Automation, Digitalization, and Artificial Intelligence in the Workplace: Implications for Political Behavior. *Annual Review of Political Science*, *25*(1), 463–484. https://doi.org/10.1146/annurev-polisci-051120-104535

Gamerdinger, A. & Willers, J. O. (2025). Solving AI ethics? Hybrid expertise and professional power in EU ethics governance. *Journal of European Public Policy*, 1–28. https://doi.org/10.1080/13501763.2025.2499113

Gasser, U. & Almeida, V. A. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*, *21*(6), 58–62. https://doi.org/10.1109/MIC.2017.4180835

Gassler, H., Polt, W. & Rammer, C. (2006). Schwerpunktsetzungen in der Forschungs- und Technologiepolitik - eine Analyse der Paradigmenwechsel seit 1945. *Österreichische Zeitschrift für Politikwissenschaft*, *35*(1), 7–23.

Geels, F. W. & Verhees, B. (2011). Cultural legitimacy and framing struggles in innovation journeys: A cultural-performative perspective and a case study of Dutch nuclear energy (1945–1986). *Technological Forecasting and Social Change*, *78*(6), 910–930. https://doi.org/10.1016/j.techfore.2010.12.004

Gehlen, A. (2004). *Die Seele im technischen Zeitalter und andere sozialpsychologische, soziologische und kulturanalytische Schriften: Gesamtausgabe*. *Gesamtausgabe / Arnold Gehlen: Bd. 6*. Klostermann. (Erstveröffentlichung 1957)

Gehlen, A. (2017). *Anthropologische und sozialpsychologische Untersuchungen*. Rowohlt Taschenbuch Verlag GmbH. (Erstveröffentlichung 1986)

Genus, A. & Stirling, A. (2018). Collingridge and the dilemma of control: Towards responsible and accountable innovation. *Research Policy*, *47*(1), 61–69. https://doi.org/10.1016/j.respol.2017.09.012

Geradin, D., Karanikioti, T. & Katsifis, D. (2021). GDPR Myopia: how a well-intended regulation ended up favouring large online platforms - the case of ad tech. *European Competition Journal*, *17*(1), 47–92. https://doi.org/10.1080/17441056.2020.1848059

Gesellschaft für deutsche Sprache e. V. (2025, 5. Dezember). *Wörter des Jahres 2025* [Pressemitteilung]. https://gfds.de/worter-des-jahres-2025/

Gesellschaft für Reaktorsicherheit. (1980). *Deutsche Risikostudie Kernkraftwerke: Eine Untersuchung zu dem durch Störfälle in Kernkraftwerken verursachten Risiko*. TÜV Rheinland.

Giovanola, B. & Tiribelli, S. (2022). Weapons of moral construction? On the value of fairness in algorithmic decision-making. *Ethics and Information Technology*, *24*(1), 1–13. https://doi.org/10.1007/s10676-022-09622-5

Giró Gràcia, X. & Sancho-Gil, J. M. (2021). Artificial Intelligence in Education: Big Data, Black Boxes, and Technological Solutionism. *Seminar.net*, *17*(2), 1–19. https://doi.org/10.7577/seminar.4281

Glaser, B. G. & Strauss, A. L. (2017). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Routledge. https://doi.org/10.4324/9780203793206 (Erstveröffentlichung 1967)

Glaser, K. (2013). *Über legitime Herrschaft: Grundlagen der Legitimitätstheorie*. Springer VS. https://doi.org/10.1007/978-3-658-00461-3

Grande, E. & Kriesi, H. (2013). Das Doppelgesicht der Politisierung. Zur Transformation politischer Konfliktstrukturen im Prozess der Globalisierung. In M. Zürn & M. Ecker-Ehrhardt (Hrsg.), *Die Politisierung der Weltpolitik: Umkämpfte internationale Institutionen* (S. 84–108). Suhrkamp.

Graß, D. (2015). Bildungsarbeit im Spiegel von Rechtfertigung und Kritik. Überlegungen zur Spezifität normativer Deutungsmuster im Feld schulischer Leistungspolitiken. In M. Dammayr, D. Graß & B. Rothmüller (Hrsg.), *Sozialtheorie. Legitimität: Gesellschaftliche, politische und wissenschaftliche Bruchlinien der Rechtfertigung* (S. 289–313). transcript-Verlag.

Green, B. (2021a). The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice. *Journal of Social Computing*, *2*(3), 209–225. https://doi.org/10.23919/JSC.2021.0018

Green, B. (2021b). Data Science as Political Action: Grounding Data Science in a Politics of Justice. *Journal of Social Computing*, *2*(3), 249–265. https://doi.org/10.23919/JSC.2021.0029

Green, B. & Viljoen, S. (2020). Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor & G. Zanfir-Fortuna (Hrsg.), *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (S. 19–31). ACM. https://doi.org/10.1145/3351095.3372840

Grimm, P. W., Grossman, M. R. & Cormack, G. V. (2021). Artificial Intelligence as Evidence. *Northwestern Journal of Technology and Intellectual Property*, *19*(1), 9–106.

Grunwald, A. (2003). Zukunftstechnologien und Demokratie: Zur Rolle der Technikfolgenabschätzung für demokratische Technikgestaltung. In K. Mensch & J. C. Schmidt (Hrsg.), *Technik und Demokratie: Zwischen Expertokratie, Parlament und Bürgerbeteiligung* (S. 197–211). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-97599-7_12

Grüttner, M. (2013). *Zur Bedeutung von „Konventionen" in der Arbeitsverwaltung: Die Vergabepraxis des Gründungszuschusses*. *IAB-Bibliothek Forschungsarbeiten: Bd. 342*. W. Bertelsmann Verlag.

Guenduez, A. A. & Mettler, T. (2023). Strategically constructed narratives on artificial intelligence: What stories are told in governmental artificial intelligence policies? *Government Information Quarterly*, *40*(1), 1–13. https://doi.org/10.1016/j.giq.2022.101719

Guest, G., MacQueen, K. M. & Namey, E. E. (2012). *Applied Thematic Analysis*. SAGE Publications.

Guihot, M., Matthew, A. F. & Suzor, N. P. (2017). Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence. *Vanderbilt Journal of Entertainment & Technology Law*, *20*(2), 385–456.

Gunn, P. (2019). Against Epistocracy. *Critical Review*, *31*(1), 26–82. https://doi.org/10.1080/08913811.2019.1609842

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S. & Yang, G.-Z. (2019). XAI-Explainable artificial intelligence. *Science Robotics*, *4*(37). https://doi.org/10.1126/scirobotics.aay7120

Gutiérrez, J. L. M. (2023). On actor-network theory and algorithms: ChatGPT and the new power relationships in the age of AI. *AI and Ethics*, 1–14. https://doi.org/10.1007/s43681-023-00314-4

Habermas, J. (1976a). Legitimationsprobleme im modernen Staat. In P. Kielmansegg (Hrsg.), *Politische Vierteljahresschrift Sonderheft: Bd. 7. Legitimationsprobleme politischer Systeme: Tagung der Deutschen Vereinigung für Politische Wissenschaft in Duisburg, Herbst 1975* (S. 39–61). Westdeutscher Verlag. https://doi.org/10.1007/978-3-322-88717-7_2

Habermas, J. (1976b). *Zur Rekonstruktion des Historischen Materialismus* (2. Aufl.). *Suhrkamp-Taschenbuch Wissenschaft: Bd. 154*. Suhrkamp.

Habermas, J. (1968). *Technik und Wissenschaft als „Ideologie"*. Suhrkamp Verlag.

Habermas, J. (1985). *Die neue Unübersichtlichkeit*. Suhrkamp Verlag.

Habermas, J. (1992). *Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Suhrkamp.

Hacker, P. (2024). Sustainable AI Regulation. *Common Market Law Review*, *61*(2), 345–386. https://doi.org/10.54648/cola2024025

Hacker, P., Engel, A. & Mauer, M. (2023). Regulating ChatGPT and other Large Generative AI Models. In *FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (S. 1112–1123). ACM. https://doi.org/10.1145/3593013.3594067

Haenlein, M. & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, *61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, *30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hampel, J., Zwick, M. & Goldschmidt, R. (2018). *Technikradar 2018: Was die Deutschen über Technik denken.* acatech - Deutsche Akademie der Technikwissenschaften; Körber-Stiftung; ZIRIUS - Zentrum für interdisziplinäre Risiko- und Innvoationsforschung, Universität Stuttgart. https://www.acatech.de/publikation/technikradar-2018-was-die-deutschen-ueber-technik-denken/

Hao, K. (20. August 2020). The UK Exam Debacle Reminds Us that Algorithms Can't Fix Broken Systems. *MIT Technology Review*. https://www.technologyreview.com/2020/08/20/1007502/uk-exam-algorithm-cant-fix-broken-system/

Harfst, P. & Wiesner, C. (2019). Politische Legitimität und Legitimation: Vergleichende Perspektiven. In C. Wiesner & P. Harfst (Hrsg.), *Legitimität und Legitimation: Vergleichende Perspektiven* (S. 1–9). Springer VS.

Hartmann, M. (2008). Rechtfertigungsordnungen und Anerkennungsordnungen. Zum Vergleich zweier Theoriemodelle. *WestEnd - Neue Zeitschrift für Sozialforschung*, *5*(2), 104–119.

Häußling, R. (2019). *Techniksoziologie: Eine Einführung* (2., überarbeitete und aktualisierte Auflage). Verlag Barbara Budrich.

Hay, C. (2007). *Why We Hate Politics*. Polity Press.

Hay, C. (2014). Depoliticisation as Process, Governance as Practice: What Did the 'First Wave' Get Wrong and Do We Need a 'Second Wave' to Put it Right? *Policy & Politics*, *42*(2), 293–311. https://doi.org/10.1332/030557314X13959960668217

Heidegger, M. (2011). *Die Technik und die Kehre* (12. Aufl.). Klett-Cotta. (Erstveröffentlichung 1962)

Henin, C. & Le Métayer, D. (2021). Beyond Explainability: Justifiability and Contestability of Algorithmic Decision Systems. *AI & SOCIETY*, *37*(4), 1397–1410. https://doi.org/10.1007/s00146-021-01251-8

Henneguelle, A. (2023). Socio-economics of Quantification and Value: The Perspective of Convention Theory. In R. Diaz-Bone & G. de Larquier (Hrsg.), *Handbook of Economics and Sociology of Conventions* (S. 1–20). Springer International Publishing. https://doi.org/10.1007/978-3-030-52130-1_55-2

Hennis, W. (1976a). Legitimität: Zu einer Kategorie der bürgerlichen Gesellschaft. In P. Kielmansegg (Hrsg.), *Politische Vierteljahresschrift Sonderheft: Bd. 7. Legitimationsprobleme politischer Systeme: Tagung der Deutschen Vereinigung für Politische Wissenschaft in Duisburg, Herbst 1975* (S. 9–38). Westdeutscher Verlag. https://doi.org/10.1007/978-3-322-88717-7_1

Hesse, W. (2020). Das Zerstörungspotenzial von Big Data und Künstlicher Intelligenz für die Demokratie. *Informatik Spektrum*, *43*(5), 339–346. https://doi.org/10.1007/s00287-020-01283-2

Heyen, N. B., Friedewald, M., Jochem, E., Reiß, T. & Thielmann, A. (2024). Assessing Technological Innovations: From Early Warning to the Governance of Socio-technical Transformations. In J. Edler & R. Walz (Hrsg.), *Systems and Innovation Research in Transition: Research Questions and Trends in Historical Perspective* (S. 145–172). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-66100-6_6

HFD (2025). *Das Hochschulforum Digitalisierung: Über uns*. https://hochschulforumdigitalisierung.de/ueber-uns/

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J. & Welch, V. A. (Hrsg.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (Second edition). John Wiley & Sons.

Hildebrandt, K. (2018). Der Mediendiskurs zu Umweltthemen: Eine vergleichende Analyse der Berichterstattung über den Emissionshandel in Deutschland und Großbritannien. In M. Lünenborg & S. Sell (Hrsg.), *Politischer Journalismus im Fokus der Journalistik* (S. 191–213). Springer VS.

Hildebrandt, M. (2016). Law as Information in the Era of Data-Driven Agency. *The Modern Law Review*, *79*(1), 1–30. https://doi.org/10.1111/1468-2230.12165

Hilgartner, S., Prainsack, B. & Hurlbut, J. B. (2017). Ethics as Governance in Genomics and Beyond. In U. Felt, R. Fouché, C. A. Miller & L. Smith-Doerr (Hrsg.), *The Handbook of Science and Technology Studies* (S. 823–852). The MIT Press.

Hirsch-Kreinsen, H. (2023). *Das Versprechen der Künstlichen Intelligenz: Gesellschaftliche Dynamik einer Schlüsseltechnologie*. Campus Verlag. https://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-3230073

Hirsch-Kreinsen, H. (2024). Artificial intelligence: a "promising technology". *AI & SOCIETY*, *39*(4), 1641–1652. https://doi.org/10.1007/s00146-023-01629-w

Hirsch-Kreinsen, H. & Krokowski, T. (2024a). Promises and Myths of Artificial Intelligence. *Weizenbaum Journal of the Digital Society*, *4*(1), 1–8. https://doi.org/10.34669/WI.WJDS/4.1.6

Hirsch-Kreinsen, H. & Krokowski, T. (2024b). Trustworthy AI: AI made in Germany and Europe? *AI & SOCIETY*, *39*(6), 2921–2931. https://doi.org/10.1007/s00146-023-01808-9

Hjaltalin, I. T. & Sigurdarson, H. T. (2024). The strategic use of AI in the public sector: A public values analysis of national AI strategies. *Government Information Quarterly*, *41*(1), 101914. https://doi.org/10.1016/j.giq.2024.101914

Hoffmann, N. (2018). *Dokumentenanalyse in der Bildungs- und Sozialforschung: Überblick und Einführung*. *Grundlagentexte Methoden*. Beltz.

Holm, E. A. (2019). In defense of the black box. *Science*, *364*(6435), 26–27. https://doi.org/10.1126/science.aax0162

Holmes, W. & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Addison-Wesley. https://permalink.obvsg.at/AC01223128

Holzinger, A., Langs, G., Denk, H., Zatloukal, K. & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *WIREs Data Mining and Knowledge Discovery*, *9*(4), Artikel e1312, 1–13. https://doi.org/10.1002/widm.1312

Honneth, A. (2008). Verflüssigung des Sozialen. Zur Gesellschaftstheorie von Luc Boltanski und Laurent Thévenot. *Zeitschrift für Rechtssoziologie*, *5*(2), 84–103.

Hood, C. (1991). A Public Management for All Seasons? *Public Administration*, *69*(1), 3–19. https://doi.org/10.1111/j.1467-9299.1991.tb00779.x

Hopf, C. & Schmidt, C. (1993). *Zum Verhältnis von innerfamilialen sozialen Erfahrungen, Persönlichkeitsentwicklung und politischen Orientierungen. Dokumentation und Erörterung des methodischen Vorgehens in einer Studie zu diesem Thema*. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-456148

Horkheimer, M. (1967). *Zur Kritik der instrumentellen Vernunft*. Fischer Taschenbuch Verlag. (Erstveröffentlichung 1947)

Horkheimer, M. & Adorno, T. W. (1969). *Dialektik der Aufklärung: Philosophische Fragmente*. S. Fischer. (Erstveröffentlichung 1947)

Horvath, K., Frei, A. I. & Steinberg, M. (2025). How do problematisations "materialise"? A neopragmatic perspective on the (in-)visibilisation of justice concerns in AIED. *British Journal of Sociology of Education*, 1–20. https://doi.org/10.1080/01425692.2025.2531262

Hughes, T. P. (1994). Technological Momentum. In M. R. Smith & L. Marx (Hrsg.), *Does Technology Drive History? The Dilemma of Technological Determinism* (S. 101–114). MIT Press.

Hutter, S. & Grande, E. (2014). Politicizing Europe in the National Electoral Arena: A Comparative Analysis of Five West European Countries, 1970-2010. *JCMS: Journal of Common Market Studies*, *52*(5), 1002–1018. https://doi.org/10.1111/jcms.12133

Ifenthaler, D. & Schumacher, C. (2016a). Learning Analytics im Hochschulkontext. *WiSt - Wirtschaftswissenschaftliches Studium*, *45*(4), 176–181. https://doi.org/10.15358/0340-1650-2016-4-176

Ifenthaler, D. & Schumacher, C. (2016b). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, *64*(5), 923–938. https://doi.org/10.1007/s11423-016-9477-y

Imdorf, C. (2011). Rechtfertigungsordnungen der schulischen Selektion. In S. K. Amos, W. Meseth & M. Proske (Hrsg.), *Öffentliche Erziehung revisited* (S. 225–245). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92615-5_11

Imdorf, C. & Leemann, R. J. (2023). Education and Conventions. In R. Diaz-Bone & G. de Larquier (Hrsg.), *Handbook of Economics and Sociology of Conventions* (S. 1–33). Springer International Publishing. https://doi.org/10.1007/978-3-030-52130-1_69-1

Janiesch, C., Zschech, P. & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jasanoff, S. S. (1987). Contested Boundaries in Policy-Relevant Science. *Social Studies of Science*, *17*(2), 195–230. https://doi.org/10.1177/030631287017002001

Jasanoff, S. S. (Hrsg.). (2004). *States of Knowledge: The Co-Production of Science and the Social Order*. Routledge.

Jasanoff, S. S. (2015). Future Imperfect: Science, Technology, and the Imaginations of Modernity. In S. S. Jasanoff & S.-H. Kim (Hrsg.), *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power* (S. 1–33). University of Chicago Press. https://doi.org/10.7208/9780226276663-001

Jasanoff, S. S. (2016). *The Ethics of Invention: Technology and the Human Future*. W. W. Norton & Company Incorporated.

Jiang, Y., Li, X., Luo, H., Yin, S. & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence*, *2*(1). https://doi.org/10.1007/s44163-022-00022-8

Jobin, A., Ienca, M. & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Johnson, J. (1998). Arguing for Deliberation: Some Skeptical Considerations. In J. Elster (Hrsg.), *Cambridge studies in the theory of democracy: Bd. 1. Deliberative Democracy* (S. 161–184). Cambridge Univ. Press.

Johnson, T., Ndlovu, L., Baiyegunhi, O. O., Lora, W. S. & Desmond, N. (2024). Coercive public health policies need context-specific ethical justifications. *Monash Bioethics Review*, 1–22. https://doi.org/10.1007/s40592-024-00218-x

Johnston, S. F. (2017). Technological Parables and Iconic Illustrations: American Technocracy and the Rhetoric of the Technological Fix. *History and Technology*, *33*(2), 196–219. https://doi.org/10.1080/07341512.2017.1336851

Jungherr, A. (2023). Artificial Intelligence and Democracy: A Conceptual Framework. *Social Media + Society*, *9*(3), 1–14. https://doi.org/10.1177/20563051231186353

Just, N. & Latzer, M. (2017). Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, *39*(2), 238–258. https://doi.org/10.1177/0163443716643157

Justo-Hanani, R. (2022). The politics of Artificial Intelligence regulation and governance reform in the European Union. *Policy Sciences*, *55*(1), 137–159. https://doi.org/10.1007/s11077-022-09452-8

Kajava, K., Gonzalez Torres, A. P., Rannisto, A. & Sakai, S. (2025). Justifying AI regulation: Examining multi-stakeholder responses to the AI Act. *Telematics and Informatics, 99*, 1–12. https://doi.org/10.1016/j.tele.2025.102278

Kalluri, P. (2020). Don‘t ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*, *583*(7815), 169. https://doi.org/10.1038/d41586-020-02003-2

Kappler, K. E. (2025). Health, Algorithms, and AI: Perspectives from Economics and Sociology of Conventions. In R. Diaz-Bone & G. de Larquier (Hrsg.), *Handbook of Economics and Sociology of Conventions* (S. 1–20). Springer Nature Switzerland. https://doi.org/10.1007/978-3-030-52130-147-1

Katz, R. S. & Mair, P. (1995). Changing Models of Party Organization and Party Democracy: The Emergence of the Cartel Party. *Party Politics*, *1*(1), 5–28. https://doi.org/10.1177/1354068895001001001

Katzenbach, C. (2018). *Die Regeln digitaler Kommunikation: Governance zwischen Norm, Diskurs und Technik*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19337-9

Katzenbach, C. & Ulbricht, L. (2019). Algorithmic Governance. *Internet Policy Review*, *8*(4), 1–18. https://doi.org/10.14763/2019.4.1424

Kaun, A. (2021). Suing the algorithm: The mundanization of automated decision-making in public services through litigation. *Information, Communication & Society*, 1–17. https://doi.org/10.1080/1369118X.2021.1924827

Keck, O. (1984a). Der Schnelle Brüter: Ein Lehrstück für die Technologiepolitik. *Politische Vierteljahresschrift*, *25*(3), 296–315.

Keck, O. (1984b). *Der Schnelle Brüter: Eine Fallstudie über Entscheidungsprozesse in der Großtechnik*. Campus-Verlag.

Keller, B., Baleis, J., Starke, C. & Marcinkowski, F. (2019). *Machine Learning and Artificial Intelligence in Higher Education: A State-of-the-Art Report on the German University Landscape*. Working Paper Series FAIR, Working Paper No. 1. https://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Sozialwissenschaften/Kommunikations-_und_Medienwissenschaft/KMW_I/Working_Paper/Keller_et_al.__2019__-_AI_in_Higher_Education.pdf

Keller, B., Lünich, M. & Marcinkowski, F. (2022). How Is Socially Responsible Academic Performance Prediction Possible? Insights From a Concept of Perceived AI Fairness. In F. Almaraz-Menéndez, A. Maz-Machado, C. López-Esteban & C. Almaraz-López (Hrsg.), *Advances in Higher Education and Professional Development. Strategy, Policy, Practice, and Governance for AI in Higher Education Institutions* (S. 126–155). IGI Global. https://doi.org/10.4018/978-1-7998-9247-2.ch006

Keller, B., Lünich, M. & Marcinkowski, F. (2025). Die Interdependenz von Learning Analytics und Studierenden: Implikationen studentischer Wahrnehmungen für eine sozialverantwortliche KI-Implementierung. In R. Queckenberg, J. Leschke & M. Persike (Hrsg.), *Learning Analytics, Artificial Intelligence und Data Mining in der Hochschulbildung: Beiträge zur Learning AID 2024* (S. 73–87). transcript.

Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T. & Woodruff, A. (2021). Exciting, Useful, Worrying, Futuristic: Public Perception of Artificial Intelligence in 8 Countries. In M. Fourcade, B. Kuipers, S. Lazar & D. Mulligan (Hrsg.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (S. 627–637). ACM. https://doi.org/10.1145/3461702.3462605

Kellner, D. (2001). Feenberg‘s Questioning Technology. *Theory, Culture & Society*, *18*(1), 155–162. https://doi.org/10.1177/02632760122051562

Kelty, C. M. (2014). The Fog of Freedom. In T. Gillespie, P. J. Boczkowski & K. A. Foot (Hrsg.), *Media Technologies: Essays on Communication, Materiality, and Society* (S. 195–220). The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0010

Kerr, A., Barry, M. & Kelleher, J. D. (2020). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. *Big Data & Society*, *7*(1), 1-12. https://doi.org/10.1177/2053951720915939

Khanal, S., Zhang, H. & Taeihagh, A. (2025). Why and how is the power of Big Tech increasing in the policy process? The case of generative AI. *Policy and Society*, *44*(1), 52–69. https://doi.org/10.1093/polsoc/puae012

Kieslich, K., Lünich, M. & Došenović, P. (2023). Ever Heard of Ethical AI? Investigating the Salience of Ethical AI Issues among the German Population. *International Journal of Human–Computer Interaction*, *40*(11), 2986–2999. https://doi.org/10.1080/10447318.2023.2178612

Kim, S. W., Kong, J. H., Lee, S. W. & Lee, S [Seungchul] (2022). Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: A Review. *International Journal of Precision Engineering and Manufacturing*, *23*(1), 111–129. https://doi.org/10.1007/s12541-021-00600-3

Kirk, J. & Miller, M. L. (1989). *Reliability and Validity in Qualitative Research*. SAGE Publications.

Kizilcec, R. F. & Lee, H. (2022). Algorithmic fairness in education. In W. Holmes & K. Porayska-Pomsta (Hrsg.), *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates* (S. 174–202). Routledge. https://doi.org/10.4324/9780429329067-10

Kleinberg, J., Ludwig, J., Mullainathan, S. & Sunstein, C. R. (2018). Discrimination in the Age of Algorithms. *Journal of Legal Analysis*, *10*, 113–174. https://doi.org/10.1093/jla/laz001

KMK (2021). *Das Bildungswesen in der Bundesrepublik Deutschland 2019/2020: Darstellung der Kompetenzen, Strukturen und bildungspolitischen Entwicklungen für den Informationsaustausch in Europa*. https://www.kmk.org/fileadmin/Dateien/pdf/Eurydice/Bildungswesen-dt-pdfs/dossier_de_ebook.pdf

Kneip, S. & Merkel, W. (2020). Demokratische Legitimität: Ein theoretisches Konztept in empirisch-analytischer Absicht. In S. Kneip, W. Merkel & B. Weßels (Hrsg.), *Legitimitätsprobleme: Zur Lage der Demokratie in Deutschland* (S. 25–55). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-29558-52

Knoll, L. (2012). *Über die Rechtfertigung wirtschaftlichen Handelns: CO2-Handel in der kommunalen Energiewirtschaft*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-18983-3

Köchling, A. & Wehner, M. C. (2020). Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, *13*(3), 795–848. https://doi.org/10.1007/s40685-020-00134-w

Komljenovic, J., Sellar, S. & Birch, K. (2025). Turning universities into data-driven organisations: seven dimensions of change. *Higher Education, 89*(5), 1369–1386. https://doi.org/10.1007/s10734-024-01277-z

König, P. D. (2018). Algorithmen und die Verwaltung sozialer Komplexität: Zur Neukonfigurierung der Idee der Selbstregierung des Volkes. *Zeitschrift für Politikwissenschaft*, *28*(3), 289–312. https://doi.org/10.1007/s41358-018-0162-z

König, P. D. & Wenzelburger, G. (2019). Why Parties Take Up Digitization in their Manifestos: An Empirical Analysis of Eight Western European Economies. *Journal of European Public Policy*, *26*(11), 1678–1695. https://doi.org/10.1080/13501763.2018.1544268

König, P. D. & Wenzelburger, G. (2020). Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics. *Government Information Quarterly*, *37*(3), 101489. https://doi.org/10.1016/j.giq.2020.101489

König, P. D. & Wenzelburger, G. (2021). When Politicization Stops Algorithms in Criminal Justice. *The British Journal of Criminology*, *61*(3), 832–851. https://doi.org/10.1093/bjc/azaa099

König, P. D. & Wenzelburger, G. (2022). Between technochauvinism and human-centrism: Can algorithms improve decision-making in democratic politics? *European Political Science*, *21*(1), 132–149. https://doi.org/10.1057/s41304-020-00298-3

König, P. D., Wurster, S. & Siewert, M. B. (2023). Sustainability challenges of artificial intelligence and Citizens' regulatory preferences. *Government Information Quarterly*, *40*(4), 1–11. https://doi.org/10.1016/j.giq.2023.101863

Koops, B.-J. (2021). The concept of function creep. *Law, Innovation and Technology*, *13*(1), 29–56. https://doi.org/10.1080/17579961.2021.1898299

Kordzadeh, N. & Ghasemaghaei, M. (2021). Algorithmic Bias: Review, Synthesis, and Future Research Directions. *European Journal of Information Systems*, 1–22. https://doi.org/10.1080/0960085X.2021.1927212

Korinek, A. & Stiglitz, J. E. (2019). Artificial Intelligence and Its Implications for Income Distribution and Unemployment. In A. Agrawal, J. Gans & A. Goldfarb (Hrsg.), *The Economics of Artificial Intelligence: An Agenda. National Bureau of Economic Research Conference Report* (349 – 390). The University of Chicago Press.

Korneeva, E., Salge, T. O., Teubner, T. & Antons, D. (2023). Tracing the legitimacy of Artificial Intelligence: A longitudinal analysis of media discourse. *Technological Forecasting and Social Change*, *192*, 1–20. https://doi.org/10.1016/j.techfore.2023.122467

Köstler, L. & Ossewaarde, R. (2022). The Making of AI Society: AI Futures Frames in German Political and Media Discourses. *AI & SOCIETY*, *37*(1), 249–263. https://doi.org/10.1007/s00146-021-01161-9

Krafft, T. D., Zweig, K. A. & König, P. D. (2022). How to regulate algorithmic decision-making: A framework of regulatory requirements for different applications. *Regulation & Governance*, *16*(1), 119–136. https://doi.org/10.1111/rego.12369

Kropp, C. & Hauser, W. (2024). *Technikradar 2024: Was die Deutschen über Technik denken.* acatech - Deutsche Akademie der Technikwissenschaften; Körber-Stiftung; ZIRIUS - Zentrum für interdisziplinäre Risiko- und Innvoationsforschung, Universität Stuttgart. https://www.acatech.de/publikation/technikradar-2024/

Krücken, G. (2017). Die Transformation von Universitäten in Wettbewerbsakteure. *Beiträge zur Hochschulforschung*, *39*(3-4), 10–29.

Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung: Grundlagentexte Methoden* (5. Auflage). *Grundlagentexte Methoden*. Beltz Juventa.

Kudina, O. & van de Poel, I. (2024). A sociotechnical system perspective on AI. *Minds and Machines*, *34*(3), 1–9. https://doi.org/10.1007/s11023-024-09680-2

Kudina, O. & Verbeek, P.-P. (2019). Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy. *Science, Technology, & Human Values*, *44*(2), 291–314. https://doi.org/10.1177/0162243918793711

Kühl, N., Schemmer, M., Goutier, M. & Satzger, G. (2022). Artificial intelligence and machine learning. *Electronic Markets*, *32*(4), 2235–2244. https://doi.org/10.1007/s12525-022-00598-0

Kukkonen, A., Stoddart, M. C. J. & Ylä-Anttila, T. (2021). Actors and justifications in media debates on Arctic climate change in Finland and Canada: A network approach. *Acta Sociologica*, *64*(1), 103–117. https://doi.org/10.1177/0001699319890902

Labadze, L., Grigolia, M. & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, *20*(1), 1–17. https://doi.org/10.1186/s41239-023-00426-1

Lafaye, C. & Thévenot, L. (1993). Une justification écologique? Conflits dans l'aménagement de la nature. *Revue Française de Sociologie*, *34*(4), 495–524. https://doi.org/10.2307/3321928

Laitinen, A. & Sahlgren, O. (2021). AI Systems and Respect for Human Autonomy. *Frontiers in artificial intelligence*, *4*, 1–14. https://doi.org/10.3389/frai.2021.705164

Lambach, D. & Oppermann, K. (2023). Narratives of Digital Sovereignty in German Political Discourse. *Governance*, *36*(3), 693–709. https://doi.org/10.1111/gove.12690

Lamnek, S. & Krell, C. (2016). *Qualitative Sozialforschung* (6., vollständig überarbeitete Aufl.). Beltz. http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1116682

Lampou, R. (2023). The Integration of Artificial Intelligence in Education: Opportunities and Challenges. *Review of Artificial Intelligence in Education*, *4*, 1-12. https://doi.org/10.37497/rev.artif.intell.educ.v4i00.15

Lane, J.-E. (2000). *New Public Management*. Routledge.

Lange, S. & Schimank, U. (2007). Zwischen Konvergenz und Pfadabhängigkeit: New Public Management in den Hochschulsystemen fünf ausgewählter OECD-Länder. In K. Holzinger, H. Jörgens & C. Knill (Hrsg.), *Transfer, Diffusion und Konvergenz von Politiken* (S. 522–548). VS Verlag für Sozialwissenschaften.

Langer, M. & König, C. J. (2023). Introducing a multi-stakeholder perspective on opacity, transparency and strategies to reduce opacity in algorithm-based human resource management. *Human Resource Management Review*, *33*(1), 1–13. https://doi.org/10.1016/j.hrmr.2021.100881

Larsson, S. (2020). On the Governance of Artificial Intelligence through Ethics Guidelines. *Asian Journal of Law and Society*, *7*(3), 437–451. https://doi.org/10.1017/als.2020.19

Latour, B. (2004). On using ANT for studying information systems: a (somewhat) Socratic dialogue,. In C. Avgerou, C. Ciborra & F. Land (Hrsg.), *The Social Study of Information and Communication Technology: Innovation, Actors, and Contexts* (S. 62–76). Oxford University Press.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Latour, B. (2006). Technik ist stabilisierte Gesellschaft. In A. Belliger & D. J. Krieger (Hrsg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie* (S. 369–398). transcript-Verlag.

Law, J. (1990). Introduction: Monsters, Machines and Sociotechnical Relations. *The Sociological Review*, *38*(1_suppl), 1–23. https://doi.org/10.1111/j.1467-954X.1990.tb03346.x

Lazar, S. (2024). Power and AI: Nature and Justification. In J. B. Bullock, Y.-C. Chen, J. Himmelreich, V. M. Hudson, A. Korinek, M. M. Young & B. Zhang (Hrsg.), *The Oxford Handbook of AI Governance* (S. 198–209). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.12

Lee, J. (2020). Is Artificial Intelligence Better Than Human Clinicians in Predicting Patient Outcomes? *Journal of medical Internet research*, *22*(8), 1-5. https://doi.org/10.2196/19918

Lee, M. K., Kim, J. T. & Lizarondo, L. (2017). A Human-Centered Approach to Algorithmic Services: Considerations for Fair and Motivating Smart Community Service Management that Allocates Donations to Non-Profit Organizations. In G. Mark, S. Fussell, C. Lampe, m. schraefel, J. P. Hourcade, C. Appert & D. Wigdor (Hrsg.), *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (S. 3365–3376). ACM. https://doi.org/10.1145/3025453.3025884

Lee, M. K., Kusbit, D., Kahng, A., Kim, J. T., Yuan, X., Chan, A., See, D., Noothigattu, R., Lee, S [Siheon], Psomas, A. & Procaccia, A. D. (2019). WeBuildAI: Participatory Framwork for Algorithmic Governance. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–35. https://doi.org/10.1145/3359283

Leemann, R. J. (2014). How Schools Deal With Expectations of Gender Equality. *Swiss Journal of Sociology*, *40*(2), 215–236.

Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M.-I., Noel-Storr, A., Rader, T., Shokraneh, F., Thomas, J. & Wieland, L. S. (2019). Searching for and Selecting Studies. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page & V. A. Welch (Hrsg.), *Cochrane Handbook for Systematic Reviews of Interventions* (Second edition, S. 67–108). John Wiley & Sons.

Lemke, N., Trein, P. & Varone, F. (2024). Defining artificial intelligence as a policy problem: A discourse network analysis from Germany. *European Policy Analysis*, *10*(2), 162–187. https://doi.org/10.1002/epa2.1203

Lepri, B., Oliver, N., Letouzé, E., Pentland, A. & Vinck, P. (2018). Fair, Transparent, and Accountable Algorithmic Decision-making Processes. *Philosophy & Technology*, *31*(4), 611–627. https://doi.org/10.1007/s13347-017-0279-x

Lessmann, S., Baesens, B., Seow, H.-V. & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Lewallen, J. (2021). Emerging technologies and problem definition uncertainty: The case of cybersecurity. *Regulation & Governance*, *15*(4), 1035–1052. https://doi.org/10.1111/rego.12341

Li, W. (2024). Quantification as social technology: Integrating studies of quantification with philosophy of technology. *Technology in Society*, *78*, 1–7. https://doi.org/10.1016/j.techsoc.2024.102635

Liebig, L., Güttel, L., Jobin, A. & Katzenbach, C. (2024). Subnational AI policy: shaping AI in a multi-level governance system. *AI & SOCIETY, 39*(3), 1477–1490. https://doi.org/10.1007/s00146-022-01561-5

Linde, H. (1972). *Sachdominanz in Sozialstrukturen*. Mohr.

Lindman, J., Makinen, J. & Kasanen, E. (2023). Big Tech's power, political corporate social responsibility and regulation. *Journal of Information Technology*, *38*(2), 144–159. https://doi.org/10.1177/02683962221113596

Lombard, M., Snyder-Duch, J. & Bracken, C. C. (2002). Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. *Human Communication Research, 28*(4), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Luckmann, T. (2001). Einige Bemerkungen zum Problem der Legitimation. In C. Bohn, H. Willems, M. Breuer & M. Schorch (Hrsg.), *Sinngeneratoren: Fremd- und Selbstthematisierung in soziologisch-historischer Perspektive; [Alois Hahn zum 60. Geburtstag]* (S. 229–245). UVK-Verlagsgesellschaft.

Luhmann, N. (1983). *Legitimation durch Verfahren*. Suhrkamp Taschenbuch Verlag. (Erstveröffentlichung 1969)

Luhmann, N. (1986). *Ökologische Kommunikation: Kann die moderne Gesellschaft sich auf ökologische Gefährdungen einstellen?* Westdeutscher Verlag.

Luhmann, N. (1990). *Die Wissenschaft der Gesellschaft*. Suhrkamp.

Luhmann, N. (1991). *Soziologie des Risikos*. De Gruyter. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=3042221

Luhmann, N. (1992). *Beobachtungen der Moderne*. Westdeutscher Verlag.

Lünich, M. (2022). *Der Glaube an Big Data: Eine Analyse gesellschaftlicher Überzeugungen von Erkenntnis- und Nutzengewinnen aus digitalen Daten*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36368-0

Lünich, M. & Keller, B. (2024). Explainable Artificial Intelligence for Academic Performance Prediction. An Experimental Study on the Impact of Accuracy and Simplicity of Decision Trees on Causability and Fairness Perceptions. In *FAccT ,24: The 2024 ACM Conference on Fairness, Accountability, and Transparency,* Rio de Janeiro, Brazil.

Lünich, M., Keller, B. & Marcinkowski, F. (2024a). Diverging perceptions of artificial intelligence in higher education: A Comparison of Student and Public Assessments on Risks and Damages of Academic Performance Prediction in Germany. *Computers and Education: Artificial Intelligence*, *7*, 1–15. https://doi.org/10.1016/j.caeai.2024.100305

Lünich, M., Keller, B. & Marcinkowski, F. (2024b). Fairness of Academic Performance Prediction for the Distribution of Support Measures for Students: Differences in Perceived Fairness of Distributive Justice Norms. *Technology, Knowledge and Learning*, *29*(2), 1079–1107. https://doi.org/10.1007/s10758-023-09698-y

Ma, L. (2020). Rethinking Democratizing Potential of Digital Technology: A Review of Technology and Communication Studies. *Journal of Information, Communication and Ethics in Society*, *18*(1), 140–156. https://doi.org/10.1108/JICES-02-2019-0022

Macgilchrist, F. (2019). Cruel optimism in edtech: when the digital data practices of educational technology providers inadvertently hinder educational equity. *Learning, Media and Technology*, *44*(1), 77–86. https://doi.org/10.1080/17439884.2018.1556217

Macgilchrist, F. (2021). What is 'critical' in critical studies of edtech? Three responses. *Learning, Media and Technology*, *46*(3), 243–249. https://doi.org/10.1080/17439884.2021.1958843

Madaio, M., Blodgett, S. L., Mayfield, E. & Dixon-Román, E. (2022). Beyond "fairness": Structural (in)justice lenses on AI for education. In W. Holmes & K. Porayska-Pomsta (Hrsg.), *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates* (S. 203–239). Routledge. https://doi.org/10.4324/9780429329067-11

Mai, M. (2011). *Technik, Wissenschaft und Politik: Studien zur Techniksoziologie und Technikgovernance*. VS Verlag für Sozialwissenschaften.

Makhlouf, K., Zhioua, S. & Palamidessi, C. (2021). Machine learning fairness notions: Bridging the gap with real-world applications. *Information Processing & Management*, *58*(5), 102642. https://doi.org/10.1016/j.ipm.2021.102642

Malik, Y. S., Sircar, S., Bhat, S., Ansari, M. I., Pande, T., Kumar, P., Mathapati, B., Balasubramanian, G., Kaushik, R., Natesan, S., Ezzikouri, S., El Zowalaty, M. E. & Dhama, K. (2021). How artificial intelligence may help the Covid-19 pandemic: Pitfalls and lessons for the future. *Reviews in medical virology*, *31*(5), 1–11. https://doi.org/10.1002/rmv.2205

Manheim, K. & Kaplan, L. (2019). Artificial Intelligence: Risks to Privacy and Democracy. *The Yale Journal of Law & Technology*, *21*(1), 106–188.

Mäntymäki, M., Minkkinen, M., Birkstedt, T. & Viljanen, M. (2022). Defining Organizational AI Governance. *AI and Ethics*, *2*(4), 603–609. https://doi.org/10.1007/s43681-022-00143-x

Marcinkowski, F. & Flaßhoff, F. G. (2024). Exploring the common wisdom on artificial intelligence and its political consequences: The case of Germany. In M. W. Bauer & B. Schiele (Hrsg.), *AI and Common Sense: Ambitions and Frictions* (S. 126–142). Routledge. https://doi.org/10.4324/9781032626192-14

Marcinkowski, F., Kieslich, K., Starke, C. & Lünich, M. (2020). Implications of AI (Un-)Fairness in Higher Education Admissions. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor & G. Zanfir-Fortuna (Hrsg.), *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (S. 122–130). ACM. https://doi.org/10.1145/3351095.3372867

Marcuse, H. (1941). Some Social Implications of Modern Technology. *Zeitschrift für Sozialforschung*, *9*(3), 414–439. https://doi.org/10.5840/zfs19419339

Marenco, M. & Seidl, T. (2021). The Discursive Construction of Digitalization: A Comparative Analysis of National Discourses on the Digital Future of Work. *European Political Science Review*, *13*(3), 391–409. https://doi.org/10.1017/S175577392100014X

Marmolejo-Ramos, F., Workman, T., Walker, C., Lenihan, D., Moulds, S., Correa, J. C., Hanea, A. M. & Sonna, B. (2022). AI-powered narrative building for facilitating public participation and engagement. *Discover Artificial Intelligence*, *2*(7), 1–16. https://doi.org/10.1007/s44163-022-00023-7

Martin, K. & Waldman, A. (2022). Are Algorithmic Decisions Legitimate? The Effect of Process and Outcomes on Perceptions of Legitimacy of AI Decisions. *Journal of Business Ethics*, 1–18. https://doi.org/10.1007/s10551-021-05032-7

Martins, B. O., Lavallée, C. & Silkoset, A. (2021). Drone Use for COVID-19 Related Problems: Techno-solutionism and its Societal Implications. *Global Policy, 12*(5), 603–612. https://doi.org/10.1111/1758-5899.13007

Marx, L. & Smith, M. R. (1994). Introduction. In M. R. Smith & L. Marx (Hrsg.), *Does Technology Drive History? The Dilemma of Technological Determinism.* MIT Press.

Mau, S. (2020). Numbers matter! The society of indicators, scores and ratings. *International Studies in Sociology of Education*, *29*(1-2), 19–37. https://doi.org/10.1080/09620214.2019.1668287

Mayer, A.-S., Strich, F. & Fiedler, M. (2020). Unintended Consequences of Introducing AI Systems for Decision Making. *MIS Quarterly Executive*, 239–257. https://doi.org/10.17705/2msqe.00036

Mayring, P. (2016). *Einführung in die qualitative Sozialforschung* (6., neu ausgestattete, überarbeitete Aufl.). Beltz. http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1127318

Mayring, P. (2019). Qualitative Content Analysis: Demarcation, Varieties, Developments. *Forum: Qualitative Social Research*, *20*(3), 1–14. https://doi.org/10.17169/fqs-20.3.3343

Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13., überarbeitete Auflage). Beltz.

McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

McConvey, K. & Guha, S. (2024). „This is not a data problem": Algorithms and Power in Public Higher Education in Canada. In F. F. Mueller, P. Kyburz, J. R. Williamson, C. Sas, M. L. Wilson, P. T. Dugas & I. Shklovski (Hrsg.), *ACM Digital Library, Proceedings of the CHI Conference on Human Factors in Computing Systems* (S. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642451

McGraw, K. M., Best, S. & Timpone, R. (1995). „What They Say or What They Do?" The Impact of Elite Explanation and Policy Outcomes on Public Opinion. *American Journal of Political Science*, *39*(1), 53–74. https://doi.org/10.2307/2111757

McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. Routledge.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. & Galstyan, A. (2022). A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, *54*(6), 1–35. https://doi.org/10.1145/3457607

Meiering, D. & Schäfer, A. (2020). Einleitung (Ent-)Politisierung – Debatten, Modelle und Befunde. In A. Schäfer & D. Meiering (Hrsg.), *Sonderband Leviathan: Bd. 35. (Ent-)Politisierung? Die demokratische Gesellschaft im 21. Jahrhundert* (S. 11–36). Nomos. https://doi.org/10.5771/9783748904076-11

Melo-López, V.-A., Basantes-Andrade, A., Gudiño-Mejía, C.-B. & Hernández-Martínez, E. (2025). The Impact of Artificial Intelligence on Inclusive Education: A Systematic Review. *Education Sciences*, *15*(5), 1–25. https://doi.org/10.3390/educsci15050539

Mensch, K. & Schmidt, J. C. (2003). Demokratische Wissenschafts- und Technikgestaltung: Eine einleitende Skizze zur Demokratiefähigkeit von Zukunftstechnologien. In K. Mensch & J. C. Schmidt (Hrsg.), *Technik und Demokratie: Zwischen Expertokratie, Parlament und Bürgerbeteiligung* (S. 11–28). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-97599-7_1

Mey, G. & Mruck, K. (Hrsg.). (2011). *Grounded Theory Reader*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-93318-4

Milan, S. (2020). Techno-solutionism and the standard human in the making of the COVID-19 pandemic. *Big Data & Society*, *7*(2), 1-7. https://doi.org/10.1177/2053951720966781

Minh, D., Wang, H. X., Li, Y. F. & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, *55*(5), 3503–3568. https://doi.org/10.1007/s10462-021-10088-y

Mintz, B. (2021). Neoliberalism and the Crisis in Higher Education: The Cost of Ideology. *The American Journal of Economics and Sociology*, *80*(1), 79–112. https://doi.org/10.1111/ajes.12370

Mitchell, S., Potash, E., Barocas, S., D'Amour, A. & Lum, K. (2021). Algorithmic Fairness: Choices, Assumptions, and Definitions. *Annual Review of Statistics and Its Application*, *8*(1), 141–163. https://doi.org/10.1146/annurev-statistics-042720-125902

Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Mohanasundari, S. K., Kalpana, M., Madhusudhan, U., Vasanthkumar, K., B, R., Singh, R., Vashishtha, N. & Bhatia, V. (2023). Can Artificial Intelligence Replace the Unique Nursing Role? *Cureus*, *15*(12), 1-6. https://doi.org/10.7759/cureus.51150

Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, *151*(4), 264-270. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Monett, D. & Lewis, C. W. P. (2018). Getting Clarity by Defining Artificial Intelligence—A Survey. In V. C. Müller (Hrsg.), *Philosophy and Theory of Artificial Intelligence 2017* (Bd. 44, S. 212–214). Springer International Publishing. https://doi.org/10.1007/978-3-319-96448-5_21

Morandín-Ahuerma, F. (2022). What is Artificial Intelligence? *International Journal of Research Publication and Reviews*, *3*(12), 1947–1951. https://doi.org/10.55248/gengpi.2022.31261

Moravcsik, A. (2002). In Defence of the 'Democratic Deficit': Reassessing Legitimacy in the European Union. *Journal of Common Market Studies*, *40*(4), 555–792. https://doi.org/00390

Moravec, J. W. & Martínez-Bravo, M. C. (2023). Global trends in disruptive technological change: social and policy implications for education. *On the Horizon: The International Journal of Learning Futures, 31*(3/4), 147–173. https://doi.org/10.1108/OTH-02-2023-0007

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. Public Affairs/Perseus Books.

Morozov, E. (2014). The Rise of Data and the Death of Politics. *The Guardian*. https://www.theguardian.com/technology/2014/jul/20/rise-of-data-death-of-politics-evgeny-morozov-algorithmic-regulation

Morozov, E. (3. Juli 2016). Beware the 'empathy-washing' of self-proclaimed caring capitalists | Evgeny Morozov | The Guardian. *The Guardian*. https://www.theguardian.com/commentisfree/2016/jul/02/beware-technology-giants-claiming-compassion-for-refugees-evgeny-morozov

Morozov, E. (2018). Wider digitales Wunschdenken. In F. Schirrmacher (Hrsg.), *Edition Suhrkamp Sonderdruck. Technologischer Totalitarismus: Eine Debatte* (Originalausgabe, 2. Auflage, S. 23–28). Suhrkamp.

Morozov, E. (2020). The Tech ‚Solutions' for Coronavirus Take the Surveillance State to the Next Level. *The Guardian*. https://www.theguardian.com/commentisfree/2020/apr/15/tech-coronavirus-surveilance-state-digital-disrupt

Mosco, V. (2004). *The Digital Sublime: Myth, Power, and Cyberspace*. The MIT Press.

Mouffe, C. (2013). *Agonistics: Thinking the World Politically.* Verso Books.

Mouffe, C. (2015). *Über das Politische: Wider die kosmopolitische Illusion.* Suhrkamp Verlag. (Erstveröffentlichung 2005)

Munn, L. (2023). The uselessness of AI ethics. *AI and Ethics*, *3*(3), 869–877. https://doi.org/10.1007/s43681-022-00209-w

Musiani, F. (2013). Governance by algorithms. *Internet Policy Review*, *2*(3), 1–8. https://doi.org/10.14763/2013.3.188

Nachtwey, O. & Seidl, T. (2021). Ideologie und der Geist des digitalen Kapitalismus. In W. Matiaske & W. Nienhüser (Hrsg.), *Ökonomie und Gesellschaft: Jahrbuch 32. Ökonomie und Ideologie* (S. 235–266). Metropolis-Verlag.

Nachtwey, O. & Seidl, T. (2023). The Solutionist Ethic and the Spirit of Digital Capitalism. *Theory, Culture & Society*, 1–22. https://doi.org/10.1177/02632764231196829

Naseem, M., Akhund, R., Arshad, H. & Ibrahim, M. T. (2020). Exploring the Potential of Artificial Intelligence and Machine Learning to Combat COVID-19 and Existing Opportunities for LMIC: A Scoping Review. *Journal of Primary Care & Community Health*, *11*, 1-11. https://doi.org/10.1177/2150132720963634

Natale, S. & Ballatore, A. (2020). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence: The International Journal of Research into New Media Technologies, 26*(1), 3–18. https://doi.org/10.1177/1354856517715164

Nemorin, S., Vlachidis, A., Ayerakwa, H. M. & Andriotis, P. (2023). AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. *Learning, Media and Technology*, *48*(1), 38–51. https://doi.org/10.1080/17439884.2022.2095568

Nguyen, D. & Hekman, E. (2024). The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation. *AI & SOCIETY*, *39*(2), 437–451. https://doi.org/10.1007/s00146-022-01511-1

Niemann, H. (2019). *The Justification of Responsibility in the UN Security Council: Practices of Normative Ordering in International Relations*. *Routledge global cooperation series*. Routledge.

Noble, D. F. (1978). Social Choice in Machine Design: The Case of Automatically Controlled Machine Tools, and a Challenge for Labor. *Politics & Society*, *8*(3-4), 313–347. https://doi.org/10.1177/003232927800800302

Nocun, K. (2018). Datenschutz unter Druck: Fehlender Wettbewerb bei sozialen Netzwerken als Risiko für den Verbraucherschutz. In A. Roßnagel, M. Friedewald & M. Hansen (Hrsg.), *DuD-Fachbeiträge. Die Fortentwicklung des Datenschutzes (*S. 39–58). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-23727-1_3

Nokkala, T. (2016). Making the Case for Policy - Persuasiveness in Higher Education, Science and Technology Policy Discourse. *European Journal of Higher Education*, *6*(2), 144–158. https://doi.org/10.1080/21568235.2015.1086664

Nordström, M. (2022). AI under great uncertainty: implications and decision strategies for public policy. *AI & society*, *37*(4), 1703–1714. https://doi.org/10.1007/s00146-021-01263-4

Novelli, C., Casolari, F., Rotolo, A., Taddeo, M. & Floridi, L. (2024). Taking AI risks seriously: a new assessment model for the AI Act. *AI & SOCIETY*, *39*(5), 2493–2497. https://doi.org/10.1007/s00146-023-01723-z

Novelli, C., Hacker, P., Morley, J., Trondal, J. & Floridi, L. (2025). A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities. *European Journal of Risk Regulation*, *16*(2), 566–590. https://doi.org/10.1017/err.2024.57

Nullmeier, F. (2010). Einleitung. In F. Nullmeier, D. Biegoń, J. Gronau, M. Nonhoff, H. Schmidtke & S. Schneider (Hrsg.), *Staatlichkeit im Wandel: Bd. 14. Prekäre Legitimitäten: Rechtfertigung von Herrschaft in der postnationalen Konstellation* (S. 9–15). Campus Verlag.

Nullmeier, F., Geis, A. & Daase, C. (2012). Der Aufstieg der Legitimitätspolitik. Rechtfertigung und Kritik politisch-ökonomischer Ordnungen. In A. Geis, F. Nullmeier & C. Daase (Hrsg.), *Der Aufstieg der Legitimitätspolitik: Rechtfertigung und Kritik politisch-ökonomischer Ordnungen* (S. 9–39). Nomos. https://doi.org/10.5771/9783845245072-9

O'Reilly, T. (2013). Open Data and Algorithmic Regulation. In B. Goldstein & L. Dyson (Hrsg.), *Beyond Transparency: Open Data and the Future of Civic Innovation* (S. 289–300). Code for America Press.

Obendiek, A. S. & Seidl, T. (2023). The (False) Promise of Solutionism: Ideational Business Power and the Construction of Epistemic Authority in Digital Security Governance. *Journal of European Public Policy*, *30*(7), 1305–1329. https://doi.org/10.1080/13501763.2023.2172060

OECD. (2019). *Artificial Intelligence in Society*. OECD Publishing. https://doi.org/10.1787/eedfee77-en

OECD. (2024). *Recommendation of the Council on OECD Legal Instruments Artificial Intelligence*. OECD/LEGAL/0449. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Oevermann, U., Allert, T., Konau, E. & Krambeck, J. (1979). Die Methodologie einer „objektiven Hermeneutik" und ihre allgemeine forschungslogische Bedeutung in den Sozialwissenschaften. In H.-G. Soeffner (Hrsg.), *Interpretative Verfahren in den Sozial- und Textwissenschaften* (S. 352–434). Metzler.

Ogburn, W. F. (1969). *Kultur und sozialer Wandel: Ausgewählte Schriften*. Luchterhand.

O'Leary, D. E. (2013). Artificial Intelligence and Big Data. *IEEE Intelligent Systems*, *28*(2), 96–99. https://doi.org/10.1109/MIS.2013.39

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Ossewaarde, M. & Gülenç, E. (2020). National Varieties of Artificial Intelligence Discourses: Myth, Utopianism, and Solutionism in West European Policy Expectations. *Computer*, *53*(11), 53–61. https://doi.org/10.1109/MC.2020.2992290

Ouchchy, L., Coin, A. & Dubljević, V. (2020). AI in the Headlines: The Portrayal of the Ethical Issues of Artificial Intelligence in the Media. *AI & SOCIETY*, *35*(4), 927–936. https://doi.org/10.1007/s00146-020-00965-5

Owen, R., Bessant, J. & Heintz, M. (Hrsg.). (2013). *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*. Wiley.

Owen, R. & Pansera, M. (2019). Responsible Innovation and Responsible Research and Innovation. In D. Simon, S. Kuhlmann, J. Stamm & W. Canzler (Hrsg.), *Handbook on Science and Public Policy* (S. 26–48). Edward Elgar Publishing.

Parycek, P., Schmid, V. & Novak, A.-S. (2024). Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations, and Framework Conditions. *Journal of the Knowledge Economy*, *15*(2), 8390–8415. https://doi.org/10.1007/s13132-023-01433-3

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Passoth, J.-H. (2008). *Technik und Gesellschaft: Sozialwissenschaftliche Techniktheorien und die Transformationen der Moderne*. VS Verlag für Sozialwissenschaften. https://doi.org/15582

Pataranutaporn, P., Danry, V., Leong, J., Punpongsanon, P., Novy, D., Maes, P. & Sra, M. (2021). AI-generated characters for supporting personalized learning and well-being. *Nature Machine Intelligence*, *3*(12), 1013–1022. https://doi.org/10.1038/s42256-021-00417-9

Păvăloaia, V.-D. & Necula, S.-C. (2023). Artificial Intelligence as a Disruptive Technology—A Systematic Literature Review. *Electronics, 12*(5), 1–37. https://doi.org/10.3390/electronics12051102

Pawelec, M. (2022). Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions. *Digital Society: Ethics, Socio-legal and Governance of Digital Technology*, *1*(2), 1–37. https://doi.org/10.1007/s44206-022-00010-6

Perry, B. & Uuk, R. (2019). AI Governance and the Policymaking Process: Key Considerations for Reducing AI Risk. *Big Data and Cognitive Computing*, *3*(2), 1–17. https://doi.org/10.3390/bdcc3020026

Peter, T. (2015). Die Gerechtigkeit der Spitze. Über die Rechtfertigung von Exellenz im deutschen Bildungsdiskurs. In M. Dammayr, D. Graß & B. Rothmüller (Hrsg.), *Sozialtheorie. Legitimität: Gesellschaftliche, politische und wissenschaftliche Bruchlinien der Rechtfertigung* (S. 177–197). transcript-Verlag.

Petkovic, D. (2023). It is Not "Accuracy vs. Explainability"—We Need Both for Trustworthy AI Systems. *IEEE Transactions on Technology and Society*, *4*(1), 46–53. https://doi.org/10.1109/TTS.2023.3239921

Petticrew, M. & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell Publishing.

Pettit, P. (2004). Depoliticizing Democracy. *Ratio Juris*, *17*(1), 52–65. https://doi.org/10.1111/j.0952-1917.2004.00254.x

Piano, S. L. (2020). Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. *Humanities and Social Sciences Communications*, *7*(1). https://doi.org/10.1057/s41599-020-0501-9

Pielke, R. A. (2007). *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge University Press.

Pinch, T. J. & Bijker, W. E. (1984). The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science*, *14*(3), 399–441. https://doi.org/10.1177/030631284014003004

Pitkin, H. F. (1967). *The Concept of Representation*. University of California Press. https://doi.org/10.1525/9780520340503

Porter, T. M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press. https://doi.org/10.1515/9780691210544

Postman, N. (1992). *Das Technopol: Die Macht der Technologien und die Entmündigung der Gesellschaft*. S. Fischer.

Prinsloo, P. (2020). Of 'black boxes' and algorithmic decision-making in (higher) education – A commentary. *Big Data & Society, 7*(1), 1-6. https://doi.org/10.1177/2053951720933994

Prinsloo, P. & Slade, S. (2015). Student Privacy Self-Management: Implications for Learning Analytics. In J. Baron, G. Lynch, N. Maziarz, P. Blikstein, A. Merceron & G. Siemens (Hrsg.), *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (S. 83–92). ACM. https://doi.org/10.1145/2723576.2723585

Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy and Society*, *40*(2), 178–193. https://doi.org/10.1080/14494035.2021.1929728

Rahm, L. (2023a). Education, automation and AI: a genealogy of alternative futures. *Learning, Media and Technology*, *48*(1), 6–24. https://doi.org/10.1080/17439884.2021.1977948

Rahm, L. (2023b). Educational imaginaries: governance at the intersection of technology and education. *Journal of Education Policy*, *38*(1), 46–68. https://doi.org/10.1080/02680939.2021.1970233

Rahm, L. & Rahm-Skågeby, J. (2023). Imaginaries and problematisations: A heuristic lens in the age of artificial intelligence in education. *British Journal of Educational Technology, 54*(5), 1147–1159. https://doi.org/10.1111/bjet.13319

Rahwan, I. (2018). Society-in-the-loop: programming the algorithmic social contract. *Ethics and Information Technology*, *20*(1), 5–14. https://doi.org/10.1007/s10676-017-9430-8

Rammert, W. (1992). Wer oder was steuert den technischen Fortschritt? Technischer Wandel zwischen Steuerung und Evolution. *Soziale Welt*, *43*(1), 7–25.

Rammert, W. (1993). *Technik aus soziologischer Perspektive: Forschungsstand · Theorieansätze · Fallbeispiele. Ein Überblick*. Westdeutscher Verlag. https://doi.org/10.1007/978-3-322-99559-9

Rammert, W. (2016). *Technik - Handeln - Wissen: Zu einer pragmatistischen Technik- und Sozialtheorie*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-11773-3

Rammert, W. (2025). *Distributed Agency & Digital Technology*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-48787-4

Rawls, J. (1971). *A Theory of Justice*. The Belknap Press of Harvard University Press.

Rawls, J. (1996). *Political Liberalism* (Paperback ed.). *The John Dewey essays in philosophy: Bd. 4*. Columbia University Press.

Reinecke, J., van Bommel, K. & Spicer, A. (2017). When Orders of Worth Clash: Negotiating Legitimacy in Situations of Moral Multiplexity. In C. Cloutier, J.-P. Gond & B. Leca (Hrsg.), *Justification, Evaluation and Critique in the Study of Organizations: Contributions from French Pragmatist Sociology* (S. 33–72). Emerald Publishing Limited.

Rességuier, A. & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society*, *7*(2), 1–5. https://doi.org/10.1177/2053951720942541

Reyes, J. I. & Meneses, J. (2024). Is artificial intelligence an opportunity for inclusive education? A case study in a fully online university. *Ubiquity Proceedings*, *4*(1), 1–6. https://doi.org/10.5334/uproc.124

Richter, V., Katzenbach, C. & Schäfer, M. S. (2023). Imaginaries of Artificial Intelligence. In S. Lindgren (Hrsg.), *Handbook of critical studies of artificial intelligence* (S. 209–223). Edward Elgar Publishing. https://doi.org/10.4337/9781803928562.00024

Richter, V., Katzenbach, C. & Zeng, J. (2025). Negotiating AI(s) futures: competing imaginaries of AI by stakeholders in the U.S., China, and Germany. *Journal of Science Communication, 24*(2), 1–20. https://doi.org/10.22323/2.24020208

Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society*, *63*, 1–15. https://doi.org/10.1016/j.techsoc.2020.101421

Ropohl, G. (1996). *Ethik und Technikbewertung*. Suhrkamp.

Ropohl, G. (2016). *Technologische Aufklärung: Beiträge zur Technikphilosophie* (3. Auflage). *Suhrkamp-Taschenbuch Wissenschaft: Bd. 971*. Suhrkamp.

Rosanvallon, P. (2011). *Democratic Legitimacy: Impartiality, Reflexivity, Proximity*. Princeton University Press. https://www.degruyter.com/document/doi/10.1515/9781400838745/html https://doi.org/10.1515/9781400838745

Rubel, A. & Jones, K. M. L. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society*, *32*(2), 143–159. https://doi.org/10.1080/01972243.2016.1130502

Rucht, D. (2013). Globalisierungskritische Proteste als Herausforderung an die internationale Politik. In M. Zürn & M. Ecker-Ehrhardt (Hrsg.), *Die Politisierung der Weltpolitik: Umkämpfte internationale Institutionen* (S. 61–83). Suhrkamp.

Russell, S., Dewey, D. & Tegmark, M. (2015). Research Priorities for Robust and Beneficial Artificial Intelligence. *AI Magazine*, *36*(4), 105–114. https://doi.org/10.1609/aimag.v36i4.2577

Ryan, M., Christodoulou, E., Antoniou, J. & Iordanou, K. (2024). An AI ethics 'David and Goliath': value conflicts between large tech companies and their employees. *AI & SOCIETY*, *39*(2), 557–572. https://doi.org/10.1007/s00146-022-01430-1

Sætra, H. S. (2020). A shallow defence of a technocracy of artificial intelligence: Examining the political harms of algorithmic governance in the domain of government. *Technology in Society*, *62*, 1–10. https://doi.org/10.1016/j.techsoc.2020.101283

Sætra, H. S. & Selinger, E. (2024). Technological Remedies for Social Problems: Defining and Demarcating Techno-Fixes and Techno-Solutionism. *Science and engineering ethics, 30*(6), 60. https://doi.org/10.1007/s11948-024-00524-x

Sahlberg, P. (2023). Trends in global education reform since the 1990s: Looking for the right way. *International Journal of Educational Development*, *98*, 1–8. https://doi.org/10.1016/j.ijedudev.2023.102748

Salais, R. (2016). Quantification and Objectivity. From Statistical Conventions to Social Conventions. *Historical Social Research*, *41*(2), 118–134. https://doi.org/10.12759/HSR.41.2016.2.118-134

Salheiser, A. (2022). Natürliche Daten: Dokumente. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 1507–1522). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-37985-8_104

Saltman, K. J. (2020). Artificial intelligence and the technological turn of public education privatization: In defence of democratic education. *London Review of Education*, *18*(2), 196–208. https://doi.org/10.14324/LRE.18.2.04

Samoili, S., López-Cobo, M., Gómez, E., Prato, G. de, Martínez-Plumed, F. & Delipetrev, B. (2020). *AI Watch: Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*. EUR 30117 EN, Publications Office of the European Union. https://doi.org/10.2760/382730

Sanderson, C., Schleiger, E., Douglas, D., Kuhnert, P. & Lu, Q. (2024). Resolving Ethics Trade-offs in Implementing Responsible AI. In *2024 IEEE Conference on Artificial Intelligence (CAI),* Singapore, Singapore.

Santori, D. (2023). *The Quantified School: Pedegogy, Subjectivity, and Metrics*. Palgrave Macmillan UK; Imprint Palgrave Macmillan. https://doi.org/10.1057/978-1-137-58385-7

Sarcinelli, U. (1989). Symbolische Politik und politische Kultur. Das Kommunikationsritual als politische Wirklichkeit. *Politische Vierteljahresschrift*, *30*(2), 292–309.

Sarcinelli, U. (1998). Legitimität. In O. Jarren, U. Sarcinelli & U. Saxer (Hrsg.), *Politische Kommunikation in der demokratischen Gesellschaft* (S. 253–267). VS Verlag für Sozialwissenschaften.

Saretzki, T. (2001). Entstehung, Verlauf und Wirkungen von Technisierungskonflikten: Die Rolle von Bürgerinitiativen, sozialen Bewegungen und politischen Parteien. In G. Simonis, R. Martinsen & T. Saretzki (Hrsg.), *Politik und Technik: Analysen zum Verhältnis von technologischem, politischem und staatlichem Wandel am Anfang des 21. Jahrhunderts* (S. 185–210). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80387-0_11

Sato, E., Shyyan, V., Chauhan, S. & Christensen, L. (2024). Putting AI in Fair: A Framework for Equity in AI-driven Learner Models and Inclusive Assessments. *Journal of Measurement and Evaluation in Education and Psychology*, *15*, 263–281. https://doi.org/10.21031/epod.1526527

Savaget, P., Chiarini, T. & Evans, S. (2018). Empowering political participation through artificial intelligence. *Science & public policy*, *46*(3), 369–380. https://doi.org/10.1093/scipol/scy064

Scharpf, F. W. (1998). Demokratische Politik in der internationalisiserten Ökonomie. In M. T. Greven (Hrsg.), *Demokratie - eine Kultur des Westens? 20. Wissenschaftlicher Kongreß der Deutschen Vereinigung für Politische Wissenschaft* (S. 81–103). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-92308-0_5

Scharpf, F. W. (1999). *Governing in Europe: Effective and Democratic?* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198295457.001.0001

Scharpf, F. W. (2004). *Legitimationskonzepte jenseits des Nationalstaats*. MPIFG Working Paper, 04/6. Max-Planck-Institut für Gesellschaftsforschung.

Schattschneider, E. E. (1957). Intensity, Visibility, Direction and Scope. *The American Political Science Review*, *51*(4), 933–942.

Schelsky, H. (1961). *Der Mensch in der wissenschaftlichen Zivilisation*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-663-02159-9

Scherer, M. U. (2016). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, *29*(2), 353–400. https://doi.org/10.2139/ssrn.2609777

Schiff, D. (2021). Out of the laboratory and into the classroom: the future of artificial intelligence in education. *AI & SOCIETY*, *36*(1), 331–348. https://doi.org/10.1007/s00146-020-01033-8

Schiff, D. (2022). Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. *International Journal of Artificial Intelligence in Education*, *32*(3), 527–563. https://doi.org/10.1007/s40593-021-00270-2

Schiff, D., Schiff, K. J. & Pierson, P. (2022). Assessing Public Value Failure in Government Adoption of Artificial Intelligence. *Public Administration*, *100*(3), 653–673. https://doi.org/10.1111/padm.12742

Schimank, U. (2005). 'New Public Management' and the Academic Profession: Reflections on the German Situation. *Minerva*, *43*(4), 361–376. https://doi.org/10.1007/s11024-005-2472-9

Schimmelfennig, F. (2021). Politicisation management in the European Union. In E. Bressanelli, C. Koop & C. Reh (Hrsg.), *Strategic Responses to Domestic Contestation: The EU Between Politicisation and Depoliticisation.* Routledge.

Schlaile, M., Urmetzer, S., Blok, V., Andersen, A., Timmermans, J., Mueller, M., Fagerberg, J. & Pyka, A. (2017). Innovation Systems for Transformations towards Sustainability? Taking the Normative Dimension Seriously. *Sustainability*, *9*(12), 1–20. https://doi.org/10.3390/su9122253

Schmelzle, C. (2012). Zum Begriff politischer Legitimität. In A. Geis, F. Nullmeier & C. Daase (Hrsg.), *Der Aufstieg der Legitimitätspolitik: Rechtfertigung und Kritik politisch-ökonomischer Ordnungen* (S. 417–436). Nomos. https://doi.org/10.5771/9783845245072-417

Schmidt, V. A. (2002a). Does Discourse Matter in the Politics of Welfare State Adjustment? *Comparative Political Studies*, *35*(2), 168–193. https://doi.org/10.1177/0010414002035002002

Schmidt, V. A. (2002b). *The Futures of European Capitalism*. Oxford University Press. https://doi.org/10.1093/0199253684.001.0001

Schmidt, V. A. (2013). Democracy and Legitimacy in the European Union Revisited: Input, Output and 'Throughput'. *Political Studies*, *61*(1), 2–22. https://doi.org/10.1111/j.1467-9248.2012.00962.x

Schmidt, V. A. & Wood, M. (2019). Conceptualizing throughput legitimacy: Procedural mechanisms of accountability, transparency, inclusiveness and openness in EU governance. *Public Administration*, *97*(4), 727–740. https://doi.org/10.1111/padm.12615

Schopmans, H. R. (2022). From Coded Bias to Existential Threat: Expert Frames and the Epistemic Politics of AI Governance. In V. Conitzer & J. Tasioulas (Hrsg.), *ACM Digital Library, Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (S. 627–640). Association for Computing Machinery. https://doi.org/10.1145/3514094.3534161

Schreiber, A. & Schreiber, I. (2024). Bridging knowledge gap: the contribution of employees' awareness of AI cyber risks comprehensive program to reducing emerging AI digital threats. *Information & Computer Security*, *32*(5), 613–635. https://doi.org/10.1108/ICS-10-2023-0199

Schubert, K. (1991). *Politikfeldanalyse: Eine Einführung*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-85147-5

Schubert, K. & Klein, M. (2021). *Das Politiklexikon: Begriffe, Fakten, Zusammenhänge* (8., vollständig überarbeitete und erweiterte Auflage). Dietz.

Schulz-Schaeffer, I. (2008). *Technik als Gegenstand der Soziologie*. The Technical University Technology Studies Workings Papers, TUTS-WP-3-2008. Technische Universität Berlin. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-12318

Schulz-Schaeffer, I. (2021). Innovation als soziale Konstruktion von Technik und Techniknutzung. In B. Blättel-Mink, I. Schulz-Schaeffer & A. Windeler (Hrsg.), *Handbuch Innovationsforschung: Sozialwissenschaftliche Perspektiven* (S. 145–162). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-17668-6_10

Selbst, A. D., boyd, d., Friedler, S. A., Venkatasubramanian, S. & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. In *FAT* ,19: Conference on Fairness, Accountability, and Transparency,* Atlanta, GA, USA.

Selin, C. (2007). Expectations and the Emergence of Nanotechnology. *Science, Technology, & Human Values*, *32*(2), 196–220. https://doi.org/10.1177/0162243906296918

Selk, V. (2011). Die Politik der Entpolitisierung als Problem der Politikwissenschaft und der Demokratie. *Zeitschrift für Politische Theorie*, *2*, 185–200.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. *Digital futures series.* John Wiley & Sons.

Selwyn, N. (2024). The Critique of Digital Education: Time for a (Post)Critical Turn. In R. Gorur, P. Landri & R. Normand (Hrsg.), *Rethinking Sociological Critique in Contemporary Education: Reflexive Dialogue and Prospective Inquiry* (S. 48–62). Routledge. https://doi.org/10.4324/9781003279457-5

Shah, H. (2018). Algorithmic accountability. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *376*(2128), 1–6. https://doi.org/10.1098/rsta.2017.0362

Shapiro, I. (1999). Enough of Deliberation: Politics is about Interests and Power. In S. Macedo (Hrsg.), *Practical and professional ethics series. Deliberative Politics: Essays on „Democracy and Disagreement"* (S. 28–38). Oxford University Press.

Sharma, H. (2022). Mapping the Global EdTech Revolution during the Pandemic: From 'Determinism' to 'Solutionism'. In E. Mogaji, V. Jain, F. Maringe & R. E. Hinson (Hrsg.), *Re-imagining Educational Futures in Developing Countries: Lessons from Global Health Crises* (S. 119–137). Springer International Publishing. https://doi.org/10.1007/978-3-030-88234-1_7

Sharon, T. (2018). When Digital Health Meets Digital Capitalism, How Many Common Goods Are at Stake? *Big Data & Society*, *5*(2), 1-12. https://doi.org/10.1177/2053951718819032

Sharon, T. (2021). Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech‘s newfound role as global health policy makers. *Ethics and Information Technology*, *23*(Suppl 1), 45–57. https://doi.org/10.1007/s10676-020-09547-x

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, 1–10. https://doi.org/10.1016/j.ijhcs.2020.102551

Shin, D. & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284. https://doi.org/10.1016/j.chb.2019.04.019

Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human–Computer Interaction*, *36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Shrestha, Y. R., Ben-Menahem, S. M. & Krogh, G. von (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, *61*(4), 66–83. https://doi.org/10.1177/0008125619862257

Siewert, M. B. & König, P. D. (2019). On digital front-runners and late-comers: Analyzing issue competition over digitization in German subnational elections. *European Political Science Review*, *11*(2), 247–265. https://doi.org/10.1017/S1755773919000109

Siffels, L. E. (2021). Beyond Privacy vs. Health: A Justification Analysis of the Contact-Tracing Apps Debate in the Netherlands. *Ethics and Information Technology*, *23*(1), 99–103. https://doi.org/10.1007/s10676-020-09555-x

Simmons, A. J. (1999). Justification and Legitimacy. *Ethics*, *109*(4), 739–771. https://doi.org/10.1086/233944

Simmons, R. (2018). *Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System*. Ohio State Legal Studies Working Paper No. 442 (Nr. 52). University of California Davis Law Review. https://doi.org/10.2139/ssrn.3156510

Sindermann, C., Yang, H., Elhai, J. D., Yang, S., Quan, L., Li, M. & Montag, C. (2022). Acceptance and Fear of Artificial Intelligence: associations with personality in a German and a Chinese sample. *Discover Psychology*, *2*(1), 1–12. https://doi.org/10.1007/s44202-022-00020-y

Slade, S. & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Slota, S. C., Fleischmann, K. R., Greenberg, S., Verma, N., Cummings, B., Li, L. & Shenefiel, C. (2023). Many hands make many fingers to point: challenges in creating accountable AI. *AI & SOCIETY*, *38*(4), 1287–1299. https://doi.org/10.1007/s00146-021-01302-0

Smallman, M. (2018). Science to the Rescue of Contingent Progress? Comparing 10 Years of Public, Expert and Policy Discourses on New and Emerging Science and Technology in the United Kingdom. *Public Understanding of Science*, *27*(6), 655–673. https://doi.org/10.1177/0963662517706452

Smith, H. (2020). Algorithmic bias: should students pay the price? *AI & SOCIETY*, *35*(4), 1077–1078. https://doi.org/10.1007/s00146-020-01054-3

Smith Ochoa, C., Gadinger, F. & Yildiz, T. (2021). Surveillance Under Dispute: Conceptualising Narrative Legitimation Politics. *European Journal of International Security*, *6*(2), 210–232. https://doi.org/10.1017/eis.2020.23

Smithers, L. E. (2020). Every Optimization Is a Policy Failure: The Catch (and Release) of Student Success. *About Campus*, *25*(5), 11–18. https://doi.org/10.1177/1086482220961129

Smuha, N. A. (2022). Pitfalls and Pathways for Trustworthy Artificial Intelligence in Education. In W. Holmes & K. Porayska-Pomsta (Hrsg.), *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates* (S. 113–145). Routledge. https://doi.org/10.4324/9780429329067-7

Sombart, W. (1902). *Der moderne Kapitalismus. Historisch-systematische Darstellung des gesamteuropäischen Wirtschaftslebens von seinen Anfängen bis zu Gegenwart*. Duncker & Humblot.

Song, Y., Weisberg, L. R., Zhang, S., Tian, X., Boyer, K. E. & Israel, M. (2024). A framework for inclusive AI learning design for diverse learners. *Computers and Education: Artificial Intelligence*, *6*, 1–13. https://doi.org/10.1016/j.caeai.2024.100212

Sørensen, E. & Ansell, C. (2023). Towards a Concept of Political Robustness. *Political Studies*, *71*(1), 69–88. https://doi.org/10.1177/0032321721999974

Spencer, D. (2017). Work in and beyond the Second Machine Age: the politics of production and digital technologies. *Work, Employment and Society*, *31*(1), 142–152. https://doi.org/10.1177/0950017016645716

Star, S. L. (1990). Power, Technology and the Phenomenology of Conventions: On being Allergic to Onions. *The Sociological Review*, *38*(1_suppl), 26–56. https://doi.org/10.1111/j.1467-954X.1990.tb03347.x

Starke, C., Baleis, J., Keller, B. & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society*, *9*(2), 1–16. https://doi.org/10.1177/20539517221115189

Starke, C. & Lünich, M. (2020). Artificial Intelligence for Political Decision-Making in the European Union: Effects on Citizens' Perceptions of Input, Throughput, and Output Legitimacy. *Data & Policy*, *2*, 1–17. https://doi.org/10.1017/dap.2020.19

Steyvers, M. & Kumar, A. (2024). Three Challenges for AI-Assisted Decision-Making. *Perspectives On Psychological Science: A Journal of the Association for Psychological Science*, *19*(5), 722–734. https://doi.org/10.1177/17456916231181102

Stilgoe, J. (2020). *Who's Driving Innovation? New Technologies and the Collaborative State*. Springer International Publishing. https://doi.org/10.1007/978-3-030-32320-2

Stilgoe, J., Owen, R. & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, *42*(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Straßenberger, G. (2020). Die Rückkehr des Politischen? Anmerkungen zu re- und destabilisierenden Effekten radikaldemokratischer Protestartikulation. In A. Schäfer & D. Meiering (Hrsg.), *Sonderband Leviathan: Bd. 35. (Ent-)Politisierung? Die demokratische Gesellschaft im 21. Jahrhundert* (S. 307–329). Nomos. https://doi.org/10.5771/9783748904076-307

Straßenberger, G. (2021). Konfliktaustragung und Konfliktbegrenzung. Zum Verhältnis von Demokratie und Recht in agonalen Politiktheorien. In C. Schmidt & B. Zabel (Hrsg.), *Politik im Rechtsstaat* (S. 43–60). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783845299938-43

Stockmann, D. (2023). Tech companies and the public interest: the role of the state in governing social media platforms. *Information, Communication & Society, 26*(1), 1–15. https://doi.org/10.1080/1369118X.2022.2032796

Suchman, M. C. (1995). Managing Legitimacy: Strategic and Institutional Approaches. *The Academy of Management Review*, *20*(3), 571–610.

Sun, S., Zhai, Y., Shen, B. & Chen, Y. (2020). Newspaper Coverage of Artificial Intelligence: A Perspective of Emerging Technologies. *Telematics and Informatics*, *53*, 101433. https://doi.org/10.1016/j.tele.2020.101433

Susen, S. (2017). Remarks on the Nature of Justification: A Socio-Pragmatic Perspective. In C. Cloutier, J.-P. Gond & B. Leca (Hrsg.), *Justification, Evaluation and Critique in the Study of Organizations: Contributions from French Pragmatist Sociology* (Bd. 52, S. 349–381). Emerald Publishing Limited. https://doi.org/10.1108/S0733-558X20170000052010

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, *40*(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

Tallberg, J., Erman, E., Furendal, M., Geith, J., Klamberg, M. & Lundgren, M. (2023). The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research. *International Studies Review*, 25(3), 1–23. https://doi.org/10.1093/isr/viad040

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2), 1-14. https://doi.org/10.1177/2053951717736335

Telford, T. (11. November 2019). Apple Card Algorithm Sparks Gender Bias Allegations Against Goldman Sachs. *The Washington Post*. https://www.washingtonpost.com/business/2019/11/11/apple-card-algorithm-sparks-gender-bias-allegations-against-goldman-sachs/

Teräs, M., Suoranta, J., Teräs, H. & Curcher, M. (2020). Post-Covid-19 Education and Education Technology 'Solutionism': a Seller's Market. *Postdigital Science and Education*, *2*(3), 863–878. https://doi.org/10.1007/s42438-020-00164-x

Thaa, W. (2013). „Stuttgart 21" – Krise oder Repolitisierung der repräsentativen Demokratie? *Politische Vierteljahresschrift*, *54*(1), 1–20.

Thelen, K. (2018). Regulating Uber: The Politics of the Platform Economy in Europe and the United States. *Perspectives on Politics*, *16*(4), 938–953. https://doi.org/10.7910/DVN/QOOFYS

Thelen, K. (2019). Transitions to the Knowledge Economy in Germany, Sweden, and the Netherlands. *Comparative Politics*, *51*(2), 295–315. https://doi.org/10.5129/001041519X15647434969821

Thévenot, L. (2016). From Social Coding to Economics of Convention: A Thirty-Year Perspective on the Analysis of Qualification and Quantification Investments. *Historical Social Research*, *41*(2), 96–117. https://doi.org/10.12759/HSR.41.2016.2.96-117

Thévenot, L., Moody, M. & Lafaye, C. (2000). Forms of Valuing Nature: Arguments and Modes of Justification in French and American Environmental Disputes. In M. Lamont & L. Thévenot (Hrsg.), *Cambridge Cultural Social Studies. Rethinking comparative cultural sociology: Repertoires of evaluation in France and the United States* (2. ed., repr, S. 229–272). Cambridge University Press. https://doi.org/10.1017/CBO9780511628108.009

Thierer, A. D., Castillo, A. & Russell, R. (2017). *Artificial Intelligence and Public Policy.* Mercatus Research, Mercatus Center at George Mason University. https://doi.org/10.2139/ssrn.3021135

Thorell, L., Klingberg, T., Herlitz, A., Olsson, A. & Ådén, U. (2023, 28. April). *Stellungnahme zum Vorschlag der Nationalen Agentur für Bildung für eine nationale Digitalisierungsstrategie für das Schulsystem 2023-2027.* Karolinska Institut. https://die-pädagogische-wende.de/wp-content/uploads/2023/07/Karolinska-Stellungnahme_2023_dt.pdf

Totschnig, W. (2019). The Problem of Superintelligence: Political, not Technological. *AI & SOCIETY*, *34*(4), 907–920. https://doi.org/10.1007/s00146-017-0753-0

Tsai, Y.-S., Rates, D., Moreno-Marcos, P. M., Muñoz-Merino, P. J., Jivet, I., Scheffel, M., Drachsler, H., Delgado Kloos, C. & Gašević, D. (2020). Learning analytics in European higher education—Trends and barriers. *Computers & Education*, *155*, 1–16. https://doi.org/10.1016/j.compedu.2020.103933

Turnbull, N., Wilson, S. & Agoston, G. (2024). Revaluing and devaluing higher education beyond neoliberalism: Elitist, productivist, and populist policy and rhetoric in a field of conflict. *European Educational Research Journal*, 23(5), 631–654. https://doi.org/10.1177/14749041241272627

Ulbricht, L. (2020). Algorithmen und Politisierung. In A. Schäfer & D. Meiering (Hrsg.), *Sonderband Leviathan: Bd. 35. (Ent-)Politisierung? Die demokratische Gesellschaft im 21. Jahrhundert* (S. 255–278). Nomos. https://doi.org/10.5771/9783748904076-253

Ulnicane, I. & Aden, A. (2023). Power and politics in framing bias in Artificial Intelligence policy. *Review of Policy Research*, *40*(5), 665–687. https://doi.org/10.1111/ropr.12567

Ulnicane, I., Knight, W., Leach, T., Stahl, B. C. & Wanjiku, W.-G. (2021). Framing governance for a contested emerging technology: insights from AI policy. *Policy and Society*, *40*(2), 158–177. https://doi.org/10.1080/14494035.2020.1855800

Ulnicane, I., Knight, W., Leach, T., Stahl, B. C. & Wanjiku, W.-G. (2022). Governance of Artificial Intelligence. In M. Tinnirello (Hrsg.), *The Global Politics of Artificial Intelligence* (S. 29–56). Chapman and Hall/CRC. https://doi.org/10.1201/9780429446726-2

Umoga, U. J., Sodiya, E. O., Ugwuanyi, E. D., Jacks, B. S., Lottu, O. A., Daraojimba, O. D. & Obaigbena, A. (2024). Exploring the potential of AI-driven optimization in enhancing network performance and efficiency. *Magna Scientia Advanced Research and Reviews*, *10*(1), 368–378. https://doi.org/10.30574/msarr.2024.10.1.0028

Uyarra, E., Ribeiro, B. & Dale-Clough, L. (2019). Exploring the normative turn in regional innovation policy: responsibility and the quest for public value. *European Planning Studies*, *27*(12), 2359–2375. https://doi.org/10.1080/09654313.2019.1609425

Vaishya, R., Javaid, M., Khan, I. H. & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes & metabolic syndrome*, *14*(4), 337–339. https://doi.org/10.1016/j.dsx.2020.04.012

van der Sloot, B. & Wagensveld, Y. (2022). Deepfakes: regulatory challenges for the synthetic society. *Computer Law & Security Review*, *46*, 1–15. https://doi.org/10.1016/j.clsr.2022.105716

van Lente, H. & Rip, A. (1998). The Rise of Membrane Technology: From Rhetorics to Social Reality. *Social Studies of Science*, *28*(2), 221–254. https://doi.org/10.1177/030631298028002002

van Lente, H., Spitters, C. & Peine, A. (2013). Comparing technological hype cycles: Towards a theory. *Technological Forecasting and Social Change*, *80*(8), 1615–1628. https://doi.org/10.1016/j.techfore.2012.12.004

Veale, M. & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, *4*(2), 1-17. https://doi.org/10.1177/2053951717743530

Veale, M. & Brass, I. (2019). Administration by Algorithm? Public Management meets Public Sector Machine Learning. In K. Yeung & M. Lodge (Hrsg.), *Algorithmic Regulation* (S. 121–149). Oxford University Press.

Veale, M., Matus, K. & Gorwa, R. (2023). AI and Global Governance: Modalities, Rationales, Tensions. *Annual Review of Law and Social Science*, *19*(1), 255–275. https://doi.org/10.1146/annurev-lawsocsci-020223-040749

Veale, M. & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, *22*(4), 97–112. https://doi.org/10.9785/cri-2021-220402

Veblen, T. (1997). *The Theory of the Leisure Class: An Economic Study of Instituions*. Project Gutenberg. (Erstveröffentlichung 1912)

Vergeer, M. (2020). Artificial Intelligence in the Dutch Press: An Analysis of Topics and Trends. *Communication Studies*, *71*(3), 373–392. https://doi.org/10.1080/10510974.2020.1733038

Vertesi, J. A. & boyd, d. (2023). The Resource Bind: System Failure and Legitimacy Threats in Sociotechnical Organizations. *Sociologica, 17*(3), 25–49. https://doi.org/10.6092/ISSN.1971-8853/18894

Viberg, O., Kizilcec, R. F., Wise, A. F., Jivet, I. & Nixon, N. (2024). Advancing equity and inclusion in educational practices with AI -powered educational decision support systems (AI - EDSS). *British Journal of Educational Technology*, *55*(5), 1974–1981. https://doi.org/10.1111/bjet.13507

Volk, C. (2013). Zwischen Entpolitisierung und Radikalisierung – Zur Theorie von Demokratie und Politik in Zeiten des Widerstands. *Politische Vierteljahresschrift*, *54*(1), 75–110.

Volti, R. & Croissant, J. (2024). *Society & Technological Change: Ninth Edition*. Waveland Press. (Erstveröffentlichung 1988)

von Essen, L. & Ossewaarde, M. (2023). Artificial Intelligence and European Identity: The European Commission's Struggle for Reconciliation. *European Politics and Society, 25*(2), 375–402. https://doi.org/10.1080/23745118.2023.2244385

von Hirschhausen, C., Gerbaulet, C., Kemfert, C., Reitz, F. & Ziehm, C. (2015). Atomausstieg geht in die nächste Phase: Stromversorgung bleibt sicher - große Herausforderungen und hohe Kosten bei Rückbau und Endlagerung. *DIW Wochenbericht, 82*(22), 523–531.

Vorberg, L. (2019). Dystopische Szenarien und alternative Zukunftsvisionen: Audiovisuelle Inszenierungsstrategien rechtspopulistischer Narrative in aktuellen deutschen und US-amerikanischen Wahlkampfspots. In C. Schicha (Hrsg.), *Wahlwerbespots zur Bundestagswahl 2017* (S. 471–489). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-24405-7_24

Waldman, A. & Martin, K. (2022). Governing Algorithmic Decisions: The Role of Decision Importance and Governance on Perceived Legitimacy of Algorithmic Decisions. *Big Data & Society*, *9*(1), 1-16. https://doi.org/10.1177/20539517221100449

Waldron, J. (1987). Theoretical Foundations of Liberalism. *The Philosophical Quarterly*, *37*(147), 127–150. https://doi.org/10.2307/2220334

Wallner, J. (2008). Legitimacy and Public Policy: Seeing Beyond Effectiveness, Efficiency, and Performance. *Policy Studies Journal*, *36*(3), 421–443. https://doi.org/10.1111/j.1541-0072.2008.00275.x

Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1–37. https://doi.org/10.2478/jagi-2019-0002

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T. & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 1–19. https://doi.org/10.1016/j.eswa.2024.124167

Warner, R. & Sloan, R. H. (2021). Making Artificial Intelligence Transparent: Fairness and the Problem of Proxy Variables. *Criminal Justice Ethics*, *40*(1), 23–39. https://doi.org/10.1080/0731129X.2021.1893932

Warren, M. E. (1999). What is Political? *Journal of Theoretical Politics*, *11*(2), 207–231. https://doi.org/10.1177/0951692899011002004

Weber, M. (1993). *Die protestantische Ethik und der „Geist" des Kapitalismus: Textausgabe auf der Grundlage der ersten Fassung von 1904/05 mit einem Verzeichnis der wichtigsten Zusätze und Veränderungen aus der zweiten Fassung von 1920*. Athenäum Hain Hanstein. (Erstveröffentlichung 1905)

Weber, M. (2011). *Staatssoziologie: Soziologie der rationalen Staatsanstalt und der modernen politischen Parteien und Parlamente* (3. Aufl.). Duncker & Humblot. https://doi.org/10.3790/978-3-428-53510-1 (Erstveröffentlichung 1956)

Wehling, P. (2021). Technikfolgenabschätzung und Wissenschaft. In S. Böschen, A. Grunwald, B.-J. Krings & C. Rösch (Hrsg.), *Technikfolgenabschätzung: Handbuch für Wissenschaft und Praxis* (S. 178–190). Nomos Verlagsgesellschaft mbH & Co. KG.

Weinberg, A. M. (1967). *Reflections on Big Science*. MIT Press.

Weinberg, L. (2022). Rethinking Fairness: An Interdisciplinary Survey of Critiques of Hegemonic ML Fairness Approaches. *Journal of Artificial Intelligence Research*, *74*, 75–109. https://doi.org/10.1613/jair.1.13196

Wenzelburger, G. & Hartmann, K. (2022). Policy Formation, Termination and the Multiple Streams Framework: The Case of Introducing and Abolishing Automated University Admission in France. *Policy Studies*, *43*(5), 1075–1095. https://doi.org/10.1080/01442872.2021.1922661

Wenzelburger, G., König, P. D., Felfeli, J. & Achtziger, A. (2022). Algorithms in the public sector. Why context matters. *Public Administration*, 1–21. https://doi.org/10.1111/padm.12901

Wernet, A. (2006). *Einführung in die Interpretationstechnik der objektiven Hermeneutik*. VS Verlag für Sozialwissenschaften.

Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, *9*(11), 39–52. https://doi.org/10.22215/timreview/1282

Weyer, J. (2008). *Techniksoziologie: Genese, Gestaltung und Steuerung sozio-technischer Systeme*. Juventa-Verlag. https://doi.org/1485

Weyerer, J. C. & Langer, P. F. (2019). Garbage In, Garbage Out: The Vicious Cycle of AI-Based Discrimination in the Public Sector. In Y.-C. Chen, F. Salem & A. Zuiderwijk (Vorsitz), *dg.o 2019: 20th Annual International Conference on Digital Government Research,* Dubai United Arab Emirates.

Whittlestone, J., Nyrup, R., Alexandrova, A. & Cave, S. (2019). The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions. In V. Conitzer, G. Hadfield & S. Vallor (Hrsg.), *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (S. 195–200). ACM. https://doi.org/10.1145/3306618.3314289

Wieringa, M. (2023). "Hey SyRI, tell me about algorithmic accountability": Lessons from a landmark case. *Data & Policy*, *5*, 1–24. https://doi.org/10.1017/dap.2022.39

Wiesner, C. & Harfst, P. (2019). Legitimität als „essentially contested concept": Konzeption und Empirie. In C. Wiesner & P. Harfst (Hrsg.), *Legitimität und Legitimation: Vergleichende Perspektiven* (S. 11–32). Springer VS.

Willems, C. (2021). *Künstliche Intelligenz im politischen Diskurs*. Tectum Verlag.

Williams, R., Cloete, R., Cobbe, J., Cottrill, C., Edwards, P., Markovic, M., Naja, I., Ryan, F., Singh, J. & Pang, W. (2022). From Transparency to Accountability of Intelligent Systems: Moving Beyond Aspirations. *Data & Policy*, *4*(e7), 1–23. https://doi.org/10.1017/dap.2021.37

Williamson, B. (2016). Political Computational Thinking: Policy Networks, Digital Governance and ‚Learning to Code'. *Critical Policy Studies*, *10*(1), 39–58. https://doi.org/10.1080/19460171.2015.1052003

Williamson, B. (2024). The Social life of AI in Education. *International Journal of Artificial Intelligence in Education*, *34*(1), 97–104. https://doi.org/10.1007/s40593-023-00342-5

Williamson, B. & Komljenovic, J. (2023). Investing in imagined digital futures: the techno-financial 'futuring' of edtech investors in higher education. *Critical Studies in Education*, *64*(3), 234–249. https://doi.org/10.1080/17508487.2022.2081587

Williamson, S. M. & Prybutok, V. (2024). Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. *Applied Sciences*, *14*(2), 675. https://doi.org/10.3390/app14020675

Winfield, A. F. T. & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *376*, 1–13. https://doi.org/10.1098/rsta.2018.0085

Winkel, M. (2025). Controlling the uncontrollable: the public discourse on artificial intelligence between the positions of social and technological determinism. *AI & SOCIETY*, *40*(3), 1947–1959. https://doi.org/10.1007/s00146-024-01979-z

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, *109*(1), 121–136.

Winner, L. (2001). *Autonomous Technology: Technics-Out-of-Control as a Theme in Political Thought* (9. Aufl.). MIT Press. (Erstveröffentlichung 1977)

Wirtz, B. W. & Müller, W. M. (2019). An integrated artificial intelligence framework for public management. *Public Management Review*, *21*(7), 1076–1100. https://doi.org/10.1080/14719037.2018.1549268

Wirtz, B. W., Weyerer, J. C. & Geyer, C. (2019). Artificial Intelligence and the Public Sector - Applications and Challenges. *International Journal of Public Administration*, *42*(7), 596–615. https://doi.org/10.1080/01900692.2018.1498103

Wirtz, B. W., Weyerer, J. C. & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*, *39*(4), 1–17. https://doi.org/10.1016/j.giq.2022.101685

Wischmann, S. & Rohde, M. (2019). Neue Möglichkeiten für die Servicerobotik durch KI. In V. Wittpahl (Hrsg.), *Künstliche Intelligenz: Technologie - Anwendung - Gesellschaft* (S. 99–121). Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-58042-4_7

Wittmann, V. & Meynhardt, T. (2025). Human-centric AI governance: what the EU public values, what it really, really values. *Government Information Quarterly*, *42*(4), 1–34. https://doi.org/10.1016/j.giq.2025.102084

Wong, P.-H. (2020). Democratizing Algorithmic Fairness. *Philosophy & Technology*, *33*(2), 225–244. https://doi.org/10.1007/s13347-019-00355-w

Xu, Z. (2022). Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities. *Applied Artificial Intelligence*, *36*(1), Artikel 1025-1045. https://doi.org/10.1080/08839514.2021.2013652

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, *12*(4), 505–523. https://doi.org/10.1111/rego.12158

Zawacki-Richter, O., Marín, V. I., Bond, M. & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zeide, E. (2017). The Structural Consequences of Big Data-Driven Education. *Big data*, *5*(2), 164–172. https://doi.org/10.1089/big.2016.0061

Zeng, J., Chan, C. & Schäfer, M. S. (2022). Contested Chinese Dreams of AI? Public discourse about Artificial intelligence on WeChat and People's Daily Online. *Information, Communication & Society*, *25*(3), 319–340. https://doi.org/10.1080/1369118X.2020.1776372

Zerbe, N. (2007). Risking Regulation, Regulating Risk: Lessons from the Transatlantic Biotech Dispute. *Review of Policy Research*, *24*(5), 407–423. https://doi.org/10.1111/j.1541-1338.2007.00292.x

Zhang, B. & Dafoe, A. (2020). U.S. Public Opinion on the Governance of Artificial Intelligence. In A. Markham, J. Powles, T. Walsh & A. L. Washington (Hrsg.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (S. 187–193). ACM. https://doi.org/10.1145/3375627.3375827

Zhang, K. & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, 1–11. https://doi.org/10.1016/j.caeai.2021.100025

Zhang, X., Chan, F. T., Yan, C. & Bose, I. (2022). Towards risk-aware artificial intelligence and machine learning systems: An overview. *Decision Support Systems*, *159*, 1–13. https://doi.org/10.1016/j.dss.2022.113800

Žliobaitė, I. (2017). Measuring discrimination in algorithmic decision making. *Data Mining and Knowledge Discovery*, *31*(4), 1060–1089. https://doi.org/10.1007/s10618-017-0506-1

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Züger, T. & Asghari, H. (2023). AI for the public. How public interest theory shifts the discourse on AI. *AI & SOCIETY, 38*(2), 815–828. https://doi.org/10.1007/s00146-022-01480-5

Zürn, M. (2013). Politisierung als Konzept der Internationalen Beziehungen. In M. Zürn & M. Ecker-Ehrhardt (Hrsg.), *Die Politisierung der Weltpolitik: Umkämpfte internationale Institutionen* (S. 7–35). Suhrkamp.

Zürn, M. (2018). *A Theory of Global Governance: Authority, Legitimacy, and Contestation*. Oxford University Press.

Zürn, M., Binder, M., Ecker-Ehrhardt, M. & Radtke, K. (2007). Politische Ordnungsbildung wider Willen. *Zeitschrift für Internationale Beziehungen*, *14*(1), 129–164.

Zweig, K. A. (2019). *Algorithmische Entscheidungen: Transparenz und Kontrolle*. *Analysen & Argumente. Digitale Gesellschaft: Nr. 338*. Konrad-Adenauer-Stiftung.

Zweig, K. A., Fischer, S. & Lischka, K. (2018). *Wo Maschinen irren können: Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung*. Arbeitspapier - Implus Algorithmenethik #4. Bertelsmann Stiftung. https://doi.org/10.11586/2018006

Anhang

A1.	Verzeichnis der Analysedokumente	298
A2.	Codebuch der Dokumentenauswahl	304
A3.	Reliabilitätsprüfung der Dokumentenauswahl	313
A4.	Codierleitfaden	314
A5.	Kategoriensystem	319
A6	Reliabilitätsprüfung der Codierung der qualitativen Inhaltsanalyse	343
A7.	Relative Häufigkeit der Valenz der Rechtfertigung nach Akteursgruppen	344
A8.	Relative Häufigkeit der Cité-Bezüge nach Akteursgruppen	345
A9.	Eidesstattliche Versicherung	346

Digitaler Anhang

B1. Dokumentenarchiv

B2. Suchtrefferliste und Dokumentenauswahlprozess

B3. Quantifizierende Auswertungen

Der digitale Anhang ist über folgenden Link abrufbar:

https://osf.io/zhdm9/overview?view_only=c64ad2199472481aa33519f1fbe61e26

A1. Verzeichnis der Analysedokumente

ID	Datum	Dokumenten-/ Vorgangstitel	Dokumententyp
BT_113a	26.06.2018	Enquete-Kommission "Künstliche Intelligenz - Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale"	Antrag (Drucksache 19/2978)
BT_113b	28.10.2020	Enquete-Kommission "Künstliche Intelligenz - Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale"	Unterrichtung (Drucksache 19/23700)
BT_113c	28.06.2018	Enquete-Kommission "Künstliche Intelligenz - Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale"	BT-Plenarprotokoll 19/42
BT_113d	05.11.2020	Enquete-Kommission "Künstliche Intelligenz - Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale"	BT-Plenarprotokoll 19/189
BT_157a	09.11.2018	Künstliche Intelligenz - Auf Grundlage europäischer Werte entwickeln und zum Wohl von Gesellschaft und Umwelt gestalten	Antrag (Drucksache 19/5667)
BT_157b	13.02.2019	Künstliche Intelligenz - Auf Grundlage europäischer Werte entwickeln und zum Wohl von Gesellschaft und Umwelt gestalten	Beschlussempfehlung und Bericht (Drucksache 19/7742)
BT_157c	30.11.2018	Künstliche Intelligenz - Auf Grundlage europäischer Werte entwickeln und zum Wohl von Gesellschaft und Umwelt gestalten	BT-Plenarprotokoll 19/69
BT_157d	15.02.2019	Künstliche Intelligenz - Auf Grundlage europäischer Werte entwickeln und zum Wohl von Gesellschaft und Umwelt gestalten	BT-Plenarprotokoll 19/81
BT_162a	14.12.2018	Schwerpunkte und stringente Umsetzung der Strategie Künstliche Intelligenz der Bundesregierung	Kleine Anfrage (BT-Drucksache 19/6536)
BT_162b	05.02.2019	Schwerpunkte und stringente Umsetzung der Strategie Künstliche Intelligenz der Bundesregierung	Antwort (Drucksache 19/7566)
BT_195b	30.10.2019	Bildung in digitaler Gesellschaft dauerhaft und angemessen fördern	Beschlussempfehlung und Bericht (Drucksache 19/7742)
BT_195c	16.05.2019	Bildung in digitaler Gesellschaft dauerhaft und angemessen fördern	BT-Plenarprotokoll 19/101
BT_199b	23.10.2019	Berufliche Bildung modernisieren, Recht auf Ausbildung umsetzen	Beschlussempfehlung und Bericht (Drucksache 19/14431)
BT_199c	27.06.2019	Berufliche Bildung modernisieren, Recht auf Ausbildung umsetzen	BT-Plenarprotokoll 19/107
BR_243a	12.09.2019	Entschließung des Bundesrates „Deutschkurse für Migrantinnen und Migranten erneuern"	Antrag (Drucksache 433/19)
BT_244a	16.10.2019	Smart Germany - Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen	Antrag (Drucksache 19/14033)
BT_244b	31.01.2020	Smart Germany - Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen	Beschlussempfehlung und Bericht Ausschuss (Drucksache 19/16957)
BT_244c	17.10.2019	Smart Germany - Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen	BT-Plenarprotokoll 19/118
BT_244d	04.03.2021	Smart Germany - Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen	BT-Plenarprotokoll 19/215
BT_251c	13.02.2020	Viertes Gesetz zur Änderung des Aufstiegsfortbildungsförderungsgesetzes	Beschlussempfehlung und Bericht (Drucksache 19/17158)
BT_251e	13.12.2019	Viertes Gesetz zur Änderung des Aufstiegsfortbildungsförderungsgesetzes	BT-Plenarprotokoll 19/135
BT_251f	14.02.2020	Viertes Gesetz zur Änderung des Aufstiegsfortbildungsförderungsgesetzes	BT-Plenarprotokoll 19/147
BT_288a	20.04.2020	Alphabetisierung durch Künstliche Intelligenz - Chance für jeden	Antrag (Drucksache 19/18604)

BT_288b	18.05.2021	Alphabetisierung durch Künstliche Intelligenz - Chance für jeden	Beschlussempfehlung und Bericht (Drucksache 19/29751)
BT_315c	11.09.2020	Ausbildung in der Krise - Perspektiven für junge Menschen sichern	BT-Plenarprotokoll 19/174
BT_406a	23.03.2021	Chancen-Aufholprogramm - Lernrückstände erheben, Corona-Nachteile für Kinder und Jugendliche verhindern	Antrag (Drucksache 19/27808)
BT_406b	14.04.2021	Chancen-Aufholprogramm - Lernrückstände erheben, Corona-Nachteile für Kinder und Jugendliche verhindern	Beschlussempfehlung und Bericht (Drucksache 19/28524)
BT_455c	22.06.2022	Siebenundzwanzigstes Gesetz zur Änderung des Bundesausbildungsförderungsgesetzes (27. BAföGÄndG)	Beschlussempfehlung und Bericht (Drucksache 20/2399)
EFI_1078	12.01.2015	Jahresgutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit Deutschlands 2015	Gutachten
EFI_1082	18.01.2019	Jahresgutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit Deutschlands 2015	Gutachten
GEW_1218	27.02.2017	„Lernen mit digitalen Medien darf kein Selbstzweck sein“	Interview
GEW_1224	05.10.2018	Lernen mit dem Mini-Rechner. Programmieren in der Schule	Mitteilung
GEW_1239	24.09.2019	Welche Kinder tragen ein besonderes Risiko? Expertengespräch des GEW-Hauptvorstandes	Mitteilung
GEW_1241	18.10.2019	Schule der Zukunft. 50 Jahre Gesamtschule	Mitteilung
GEW_1242	21.10.2019	Primat des Pädagogischen. Länderserie Digitalisierung	Mitteilung
GEW_1244	01.11.2019	Learning Analytics und Big Data in der Bildung. Zur notwendigen Entwicklung eines datenpolitischen Alternativprogramms	Broschüre
GEW_1250	03.12.2019	„Milla“ ist kein Konzept für die Zukunft der Weiterbildung. GEW kritisiert E-Learning-Portal der CDU	Mitteilung
GEW_1256	07.05.2020	Keine Chance für Einstein. Coronakrise	Mitteilung
GEW_1257	07.05.2020	Learning by Doing. Covid-19	Interview
GEW_1260	29.05.2020	Schulcloud des Hasso-Platner-Instituts auf dem Vormarsch. Deutsche Auslandsschulen	Mitteilung
GEW_1265	06.07.2020	Schöne neue Digitalwelt? Pädagogik und Digitalisierung	Mitteilung
GEW_1268	22.09.2020	Bildungsgewerkschaften protestieren mit Erfolg. Abschlussprüfungen in Großbritannien	Mitteilung
GEW_1269	01.10.2020	„Das ist Dressur statt Bildung“. Verein Digitalcourage vergibt Big Brother Awards	Mitteilung
GEW_1272	06.04.2021	Kein "Weiter so wie bisher"! Coronapandemie - eine Zwischenbilanz	Mitteilung
GEW_1275	08.09.2021	Zwischen Schuldenbremse und Vermögensteuer. Bundestagswahl 2021	Mitteilung
GEW_1285	o. J.	Augmented und Virtual Reality in der Bildung	Mitteilung
GEW_1287	12.07.2022	Digitale Grundrechte	Mitteilung
GEW_1314	12.12.2022	Big Data in der Bildung	Nachricht
GEW_1338	01.05.2021	GEW-Bundesforum „Bildung in der digitalen Welt“. Projektbericht	Bericht
GWK_1038	2015	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2013	Bericht
GWK_1040	2016	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2014	Bericht
GWK_1042	2017	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2015	Bericht
GWK_1044	2018	Pakt für Forschung und Innovation. Monitoring-Bericht 2018	Bericht
GWK_1045	2019	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2017	Bericht
GWK_1047	2019	Pakt für Forschung und Innovation. Monitoring-Bericht 2019. Band II: Bericht der Wissenschaftsorganisationen	Bericht
GWK_1048	2019	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2016	Bericht
GWK_1049	2020	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2018	Bericht
GWK_1051	2020	Pakt für Forschung und Innovation. Monitoring-Bericht 2020. Teil II: Berichte der Wissenschaftsorganisationen	Bericht

GWK_1052	10.12.2020	Bund-Länder-Vereinbarung gemäß Artikel 91b Absatz 1 des Grundgesetzes über die Förderinitiative „Künstliche Intelligenz in der Hochschulbildung“	Mitteilung
GWK_1053	2021	Hochschulpakt 2020. Bericht zur Umsetzung im Jahr 2019	Bericht
GWK_1056	01.04.2021	Pakt für Forschung und Innovation Monitoring-Bericht 2021. Band	Bericht
GWK_1058	01.12.2021	Bund-Länder-Initiative zur Förderung von Künstlicher Intelligenz	Pressemitteilung
GWK_1060	01.04.2022	Pakt für Forschung und Innovation Monitoring-Bericht 2021. Band III	Bericht
GWK_1061	25.01.2022	Grundlagen der GWK 2022	Positionspapier
HFD_1187	01.03.2015	Digitales Prüfen und Bewerten im Hochschulbereich	Arbeitspapier
HFD_1188	14.03.2016	Deep Learning - Lernen mit Hilfe künstlicher Intelligenz	Blogbeitrag
HFD_1189	04.05.2016	Learning Analytics in der Hochschullehre	Blogbeitrag
HFD_1190	09.05.2016	Whitepaper: Mit Learning Analytics zu mehr Qualität in der Hochschullehre	Blogbeitrag
HFD_1191	01.11.2016	Stärkere Individualisierung der Lehre durch neue Medien	Arbeitspapier
HFD_1192	20.11.2016	Stärkere Individualisierung der Lehre durch neue Medien	Blogbeitrag
HFD_1193	25.04.2017	Eine Vision für die Zukunft digitaler Bildung	Blogbeitrag
HFD_1194	05.07.2017	"Universitäten sind nicht besonders flexibel - das ist ihre Stärke und Schwäche"	Blogbeitrag
HFD_1196	20.04.2018	Smart Learning Platform 4.0	Blogbeitrag
HFD_1200	25.06.2020	Die unerträgliche Leichtigkeit des (wissenschaftlichen) Schreibens. Mit Ghostwritern und Künstlicher Intelligenz auf der Überholspur	Blogbeitrag
HFD_1201	18.01.2021	Kann KI die Hochschulbildung revolutionieren? Ein Interview mit Claudia de Witt	Blogbeitrag
HFD_1202	01.04.2021	Künstliche Intelligenz an den Hochschulen	Arbeitspapier
HFD_1206	20.04.2021	Die Algorithmisierung des Lernens: Learning Analytics vs. Selbstbestimmtes Studium?	Blogbeitrag
HFD_1210	20.05.2021	Lernplattformen in der Hochschullehre: Lassen sich Lehrende von Learning Analytics beeinflussen?	Blogbeitrag
HFD_1211	01.10.2021	KI, Big Data und Algorithmen	Diskussionspapier
HFD_1214	21.04.2021	Von der Hochschule in die Praxis - Gespräche zu KI in der medizinischen Ausbildung	Blogbeitrag
HFD_1317	19.12.2022	ChatGPT ist erst der Anfang	Blogbeitrag
HFD_1318	19.12.2022	AI Transforming Higher Education? How a Chatbot might bring the change we've all been waiting for	Gastartikel
KMK_1065	07.12.2017	Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz	Broschüre
KMK_1076	09.12.2021	Lehren und Lernen in der digitalen Welt. Die ergänzende Empfehlung zur Strategie "Bildung in der digitalen Welt"	Strategie
LT_NI_776	11.11.2020	Sichere Gestaltung der digitalen Transformation der Wirtschaft ? welche Rolle kann die Kooperation zwischen CISPA und Universität Hannover spielen?	Antwort
LT-BE_574	07.05.2019	Entlastung der Berliner Verwaltung durch Roboter	Anfrage (Drucksache 18/18 843)
LT-BE_774	10.11.2020	Bund-Länder-Vereinbarung zur Förderung von KI in der Hochschulbildung	Vorlage (Drucksache 18/3257)
LT-BE_886	13.07.2021	Künstliche Intelligenz in der Berliner Verwaltung	Antwort (Drucksache 18/28 037)
LT-BW_793	16.12.2020	Bedeutung der Robotik für das Innovationsland Baden-Württemberg	Antrag (Drucksache 16/9558)
LT-BW_877	28.06.2021	Verordnungsvorschlag zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) COM(2021) 206 final (BR 488/21)	Mitteilung (Drucksache 17/374)
LT-BY_993	03.05.2022	Bildung in Pandemiezeiten	Schriftliche Anfrage (Drucksache 18/21208)
LT-HB_900	23.09.2021	Die Zukunft in die Schulen holen - Bedingungen für eine zeitgemäße Bildung schaffen!	Antrag (Drucksache 20/1110)

LT-HE_715	27.04.2020	Funktionen und Ausgestaltung des Schulportals	Antwort (Drucksache 20/2553)
LT-HE_765	28.10.2020	Digitale Lehre an hessischen Hochschulen	Antwort (Drucksache 20/3927)
LT-HE_979	14.03.2022	Einsatz von Avataren an hessischen Schulen	Kleine Anfrage (Drucksache 20/8068)
LT-HE_1009	01.06.2022	Frage nach Maßnahmen und Programmen zur Förderung und Unterstützung von KI-Projekten - Antwort Plenarprotokoll	Plenarprotokoll
LT-HE_1011	02.06.2022	Bildung ist in Zukunft digital: digitales Lernen an hessischen Schulen umfassend umsetzen	Plenarprotokoll
LT-HE_1027	20.07.2022	Einsatz von Avataren an hessischen Schulen	Antwort (Drucksache 20/8068)
LT-HH_497	16.07.2018	Künstliche Intelligenz in Hamburg ? Für Rot-Grün eine Spielerei oder seriöse Zukunftsplanung?	Kleine Anfrage (Drucksache 21/13799)
LT-HH_930	01.12.2021	Hamburg muss KI in der Hochschulbildung fördern	Antrag (Drucksache 22/6614)
LT-MV_779	12.11.2020	Bund-Länder-Vereinbarung gemäß Artikel 91b des Grundgesetzes über die Förderinitiative "Künstliche Intelligenz in der Hochschulbildung"	Unterrichtung (Drucksache 7/5540)
LT-NI_541	19.02.2019	Bildung in der digitalisierten Welt	Entschließungsantrag (Drucksache 18/2898)
LT-NI_547	01.03.2019	Bildung in der digitalisierten Welt	Plenarprotokoll
LT-NI_549	05.03.2019	Schul-Cloud und digitales Lernen in Niedersachsen: Wie sollen der Schutz der Schülerinnen und Schüler vor interessengeleiteter Beeinflussung und der Datenschutz gewährleistet werden?	Kleine Anfrage (Drucksache 18/3110)
LT-NI_568	10.04.2019	Schul-Cloud und digitales Lernen in Niedersachsen: Wie sollen der Schutz der Schülerinnen und Schüler vor interessengeleiteter Beeinflussung und der Datenschutz gewährleistet werden?	Antwort (Drucksache 18/3471)
LT-NI_660	13.11.2019	Chancen des "DigitalPakt Bildung" und des Grundgesetzes für die Bildung nutzen	Beschlussempfehlung (Drucksache 18/5089)
LT-NI_662	13.11.2019	Bildung in der digitalisierten Welt	Beschlussempfehlung (Drucksache 18/5089)
LT-NI_664	20.11.2019	Chancen des "DigitalPakt Bildung" und des Grundgesetzes für die Bildung nutzen	Plenarprotokoll (S. 5798-5806)
LT-NI_666	20.11.2019	Bildung in der digitalisierten Welt	Unterrichtung (Drucksache 18/5164)
LT-NI_735	15.07.2020	Bildung in der digitalisierten Welt	Antwort (Drucksache 18/7058)
LT-NI_798	19.01.2021	Sachstand Masterplan Digitalisierung	Antwort (Drucksache 18/8335)
LT-NW_753	29.09.2020	Einsetzung einer Enquetekommission "Digitale Transformation der Arbeitswelt in Nordrhein-Westfalen"	Abschlussbericht (Drucksache 17/11130)
LT-NW_773	10.11.2020	Bericht an den Landtag zur bevorstehenden Unterzeichnung von zwei Vereinbarungen zwischen Bund und Ländern nach Artikel 91b des Grundgesetzes	Verwaltungsvereinbarung
LT-NW_895	27.08.2021	Digital First oder NRW Second - Wie ist die schwarz-gelbe Digitalisierungsbilanz?	Antwort (Drucksache 17/15002)
LT-RP_464	29.11.2016	Eröffnung des "living lab smart office space" am 25. November 2016 im DFKI in Kaiserslautern	Ausschussprotokoll
LT-RP_522	15.01.2019	Künstliche Intelligenz I	Antwort (Drucksache 17/8147)
LT-RP_685	09.01.2020	Gesamtplan 2020 für IT und Digitalisierung; Schwerpunkte, Einsatz künstlicher Intelligenz, Erzeugung barrierefreier Dokumente, Übertragung weiterer Aufgaben an Dataport; sprachliche Verständlichkeit des Gesamtplans	Ausschussprotokoll
LT-RP_790	08.12.2020	Bericht der Landesregierung zur Bund-Länder-Vereinbarung "Künstliche Intelligenz (KI) in der Hochschulbildung"	Ausschussprotokoll
LT-RP_933	02.12.2021	Europa im digitalen Zeitalter: EU-Kommission setzt Rahmen für KI	Ausschussprotokoll
LT-SH_592	25.06.2019	Gesetz über die Errichtung eines Sondervermögens zur Förderung des Einsatzes von Künstlicher Intelligenz in Schleswig-Holstein und zur Änderung des Haushaltsgesetzes 2019	Gesetzentwurf (Drucksache 19/1563)

LT-SH_670	05.12.2019	Gesamtplan 2020 für IT und Digitalisierung; Schwerpunkte, Einsatz künstlicher Intelligenz, Erzeugung barrierefreier Dokumente, Übertragung weiterer Aufgaben an Dataport; sprachliche Verständlichkeit des Gesamtplans	Umdruck
LT-SH_787	02.12.2020	Unterrichtung nach dem Parlamentsinformationsgesetz zur Bund-Länder-Vereinbarung "Künstliche Intelligenz in der Hochschulbildung" gem. Art. 91 b Abs. 1 Grundgesetz; Gemeinsame Wissenschaftskonferenz, Beschlussfassung	Unterrichtung
LT-SH_830	25.03.2021	Bericht zum KI-Handlungsrahmen	Plenarprotokoll
LT-SH_897	08.09.2021	Gesetz zur Förderung der Digitalisierung und Bereitstellung von offenen Daten und zur Ermöglichung des Einsatzes von datengetriebenen Informationstechnologien in der Verwaltung (Digitalisierungsgesetz)	Gesetzentwurf (Drucksache 19/3267)
LT-SH_929	30.11.2021	Anhörung von AlgorithmWatch zum Entwurf eines Digitalisierungsgesetzes	Umdruck (Drucksache 19/3267)
LT-SH_934	09.12.2021	Fortschreibung der KI-Strategie "Künstliche Intelligenz - strategische Ziele und Handlungsfelder für SH"	Unterrichtung
LT-SH_970	23.02.2022	Gesetz zur Förderung der Digitalisierung und Bereitstellung von offenen Daten und zur Ermöglichung des Einsatzes von datengetriebenen Informationstechnologien in der Verwaltung (Digitalisierungsgesetz)	Plenarprotokoll
LT-SN_894	26.08.2021	Test von KI an sächsischen Schulen	Antwort
LT-SN_904	30.09.2021	Künstliche Intelligenz - Sachsens Potential für eine erfolgreiche Zukunft nutzen	Plenarprotokoll
LT-SN_906	07.10.2021	Künstliche Intelligenz und Datenschutz	Antwort
LT-SN_995	09.05.2022	ITS Area9 Rhapsode - Auswertung und Schlussfolgerungen	Kleine Anfrage (Drucksache 7/9846)
LT-SN_1014	07.06.2022	Unternehmen Area9 Lyceum und Testung des ITS Area Rhapsode an Schulen - Antwort	Antwort
LT-TH_733	01.07.2020	Einsetzung einer Enquetekommission "Zukunftsfähige öffentliche Infrastruktur durch Digitalisierung und Künstliche Intelligenz im Freistaat Thüringen sicherstellen: Bestandsaufnahme, Chancen und Herausforderungen"	Antrag (Drucksache 7/1129)
LT-TH_736	17.07.2020	Einsetzung einer Enquetekommission "Zukunftsfähige öffentliche Infrastruktur durch Digitalisierung und Künstliche Intelligenz im Freistaat Thüringen sicherstellen: Bestandsaufnahme, Chancen und Herausforderungen"	Plenarprotokoll
LT-TH_836	07.04.2021	Stand der Umsetzung der Thüringer Strategie für die Digitale Gesellschaft	Antwort (Drucksache 7/3024)
LT-TH_850	07.05.2021	Chancen der Digitalisierung im Gesundheitssystem nutzen: Digitalisierung der medizinischen Aus- und Weiterbildung	Plenarprotokoll
LT-TH_1310	24.08.2022	Künstliche Intelligenz im Geschäftsbereich der Landesregierung - Teil I Thüringen - Weltzien, Philipp, DIE LINKE	Kleine Anfrage
TAB_1324	05.10.2022	Künstliche Intelligenz und Distributed-Ledger-Technologien in der öffentlichen Verwaltung. TAB-Fokus Nr. 39	Arbeitsbericht
TAB_1325	05.10.2022	Künstliche Intelligenz und Distributed-Ledger-Technologien in der öffentlichen Verwaltung. Ein Überblick von Chancen und Risiken einschließlich der Darstellung international einschlägiger Praxisbeispiele	Forschungsbericht
TAB_1327	05.08.2021	Emotionserkennung mittels künstlicher Intelligenz – Perspektiven und Grenzen von Technologien zur Analyse von Gesichtsbewegungen. Themenkurzprofil Nr. 48	Forschungsbericht
TAB_1333	06.08.2021	Learning Analytics – Potenzial von KI-Systemen für Lehrende und Lernende. Themenkurzprofil Nr. 42	Forschungsbericht
TAB_1334	21.04.2021	Digitaler Wandel der Bildungs- und Arbeitswelten. TAB-Brief Nr. 47	Zeitschriftenausgabe
TAB_1335	03.01.2020	Digitale Medien in der Bildung: Endbericht zum TA-Projekt	Forschungsbericht
TAB_1336	06.08.2021	Digitale Medien in der Bildung. TAB-Fokus Nr. 12	Arbeitsbericht
WR_1093	22.01.2016	Empfehlungen zur Weiterentwicklung der MINT-Bereiche an den Hochschulen des Landes Hamburg	Empfehlung (Drs. 5085-16)
WR_1094	15.04.2016	Stellungnahme zum Neuordnungskonzept der Hochschulregion Lausitz	Stellungnahme (Drs. 5257-16)
WR_1113	21.09.2018	Neustrukturierung des Medizinstudiums und Änderung der Approbationsordnung für Ärzte | Empfehlungen der Expertenkommission zum Masterplan Medizinstudium 2020	Empfehlungen (Drs. 7271-18)
WR_1127	25.10.2019	Stellungnahme zur Weiterentwicklung der Universitätsmedizin der Westfälischen Wilhelms-Universität Münster	Stellungnahme (Drs. 8046-19)
WR_1130	25.10.2019	Stellungnahme zum Modellprojekt „Medizin neu denken“ der Universitäten Bonn und Siegen	Stellungnahme (Drs. 8047-19)
WR_1132	25.10.2019	Stellungnahme zur Weiterentwicklung der Universitätsmedizin der Universität zu Köln	Stellungnahme (Drs. 8045-19)
WR_1140	24.04.2020	Empfehlungen zur Förderung von Forschungsbauten (Förderphase 2021)	Empfehlungen (Drs. 8378-20)

WR_1143	01.10.2020	Zum Wandel in den Wissenschaften durch datenintensive Forschung	Empfehlungen (Drs. 8667-20)
WR_1158	09.07.2021	Stellungnahme zur Institutionellen Reakkreditierung der Hochschule Fresenius	Stellungnahme (Drs. 9106-21)
WR_1159	01.10.2021	Wissenschaftskommunikation	Positionspapier (Drs. 9367-21)
WR_1166	2022	Probleme und Perspektiven des Hochschulbaus 2030	Positionspapier (Drs. 9470-22)
WR_1169	21.01.2022	Stellungnahme zur Institutionellen Reakkreditierung der Hochschule Fresenius	Stellungnahme (Drs. 9474-22)
WR_1170	21.01.2022	Stellungnahme zum Antrag auf strategische Erweiterung von DIPF	Stellungnahme (Drs. 8821-21)
WR_1174	29.04.2022	Empfehlungen für eine zukunftsfähige Ausgestaltung von Studium und Lehre	Empfehlungen (Drs. 9699-22)
WR_1175	01.07.2022	Digitalisierung und Datennutzung für Gesundheitsforschung und Versorgung – Positionen und Empfehlungen	Positionspapier (Drs. 9825-22)
WR_1177	08.07.2022	Empfehlungen zur Digitalisierung in Lehre und Studium	Empfehlungen (Drs. 9848-22)
WR_1181	08.07.2022	Stellungnahme zum Antrag auf strategische Erweiterung der Senckenberg Gesellschaft für Naturforschung (SGN), Frankfurt am Main, großer strategischer Sondertatbestand im Rahmen der Ausführungsvereinbarung WGL	Stellungnahme (Drs. 9782-22)

Anmerkung: Die finale Auswahl der Untersuchungseinheiten der qualitativen Inhaltsanalyse umfasst n = 161 Dokumente. Bundestagsdokumente werden neben der Ziffern-ID zusätzlich über kombinierte Buchstaben eindeutig zugeordnet. Dieser Zusatz liegt in der im Rahmen der Dokumentenrecherche und -auswahl vorgenommenen Einschränkung auf Vorgänge begründet. Diese Vorgänge erhielten je eine ID, umfassen jedoch mehrere Dokumente, sodass die Buchstaben die einzelnen Dokumente eines Vorgangs unterscheidbar machen. | Die Zugriffslinks der Dokumente können über den digitale Anhang B2 abgerufen werden. Ein Archiv der Dokumente befindet sich ebenfalls im digitalen Anhang B1 der Arbeit.

A2. Codebuch zur Dokumentenauswahl

Handreichung zur systematischen Dokumentenrecherche

im Rahmen des Dissertationsprojekts von Birte Keller

Über die Legitimation der Einführung künstlich intelligenter Systeme

Untersuchungsziel

Zur Analyse der Rechtfertigungen des Einsatzes von künstlicher Intelligenz (KI) im Bildungssektor in Deutschland wird eine Dokumentenanalyse des bildungspolitischen Diskurses durchgeführt. In den Blick genommen werden dabei sowohl Institutionen innerhalb als auch außerhalb des politischen Systems. Die der Analyse zugrunde liegende Forschungsfrage lautet:

> *Welche Rechtfertigungen äußern bildungspolitische Akteur:innen gegenüber dem Einsatz von künstlich intelligenten Systemen im deutschen Bildungssystem?*

Die Dokumentenrecherche und -auswahl wird in fünf Schritten durchgeführt: Zuerst werden Datenbanken und Webseiten der Institutionen systematisch hinsichtlich festgelegter Suchbegriffe, die sich aus der Forschungsfrage des Vorhabens ergeben, gesichtet. Parallel dazu erfolgt in einem zweiten Schritt die Dokumentation der auf diese Weise identifizierten Dokumente. Im dritten Schritt folgt dann die Dokumentenauswahl auf der Basis der Titel, der Abstracts und ggf. einer Schlagwortsuche. Schließlich bilden der Download und die Archivierung der im vorherigen Schritt identifizierten Dokumente die letzten beiden Schritte des Recherche- und Auswahlprozesses.

Untersuchungszeitraum

Der Untersuchungszeitraum – hier gemeint als der Zeitraum, in welchem die erfassten Dokumente veröffentlicht wurden – beginnt ab dem 01. Januar 2015 und endet mit dem 31. Dezember 2022.[1] Zwar beginnt die Entwicklung von „Artificial Intelligence in Education" (AIED) und Learning-Analytics-Anwendungen bereits ab den frühen 2010er Jahren (Ferguson et al., 2016), konzentriert sich zu diesem Zeitpunkt jedoch vornehmlich auf den englischsprachigen Raum. Im deutschen Bildungssektor schreitet die Entwicklung hingegen langsamer voran (Ifenthaler & Schumacher, 2016). Das gilt sowohl für die wissenschaftliche als auch für die politische Aufmerksamkeit, worauf sowohl die Bekanntmachung der ersten Richtlinie zur Förderung von Forschung zur digitalen Hochschulbildung im Jahr 2016 (BMBF, 2016) als auch eine Diskursanalyse des Begriffs „Künstliche Intelligenz" im Deutschen Bundestag hindeuten (Willems, 2021).

Untersuchungseinheit

Untersuchungseinheiten der Analyse sind deutschsprachige Dokumente, die die Einführung und/oder Anwendung von künstlich intelligenten Systemen im Bildungssektor thematisieren. Um ein möglichst breites Spektrum an bildungspolitischen Akteur:innen abzubilden, werden dabei alle deutschen Institutionen oder Organisationen ausgewählt, die entweder potenzielle Entscheidungsträger wichtiger Policy-Entscheidungen im deutschen Bildungssystem sind oder mindestens einen bedeutenden Einfluss auf entsprechende Entscheidungen nehmen können. Berücksichtigt werden unter anderem Parlamentsmaterialien des Bundestags und der Landtage, aber auch Publikationen, Pressemitteilungen oder Blogbeiträge bildungspolitischer Institutionen, die zum oben genannten Thema zwischen 2015 und dem Stichtag der Analyse – dem 31. Dezember 2022 – veröffentlicht wurden. Konkret sollen zur Beantwortung der Forschungsfrage

[1] In einem ersten Erhebungsschritt wurde eine Suchtrefferliste für alle erschienen Dokumente im Zeitraum vom 01.01.2015 – 15.08.2022 erstellt. In einem zweiten Erhebungsschritt wurden die Dokumente für den Zeitraum vom 16.08.2022 – 31.12.2022 nacherhoben.

die öffentlich zugänglichen Dokumente der nachfolgend aufgeführten Institutionen analysiert werden:

- Deutscher Bundestag (BT)
- Deutscher Bundesrat (BR)
- Büro für Technikfolgenabschätzung (TAB)
- Deutsche Landtage (LT)
- Gemeinsame Wissenschaftskonferenz (GWK)
- Ständige Konferenz der Kultusminister (KMK)
- Wissenschaftsrat (WR)
- Expertenkommission für Forschung und Innovation (EFI)
- Hochschulrektorenkonferenz (HRK)
- Hochschulforum Digitalisierung (HFD)
- Gewerkschaft Erziehung und Wissenschaft (GEW)
- Hochschullehrerbund (HLB)
- Deutscher Lehrerverband (DL)

Datenbanken

Für den Bundestag und den Bundesrat wird dabei auf das gemeinsame „Dokumentations- und Informationssystem für Parlamentsmaterialien (DIP)“ zurückgegriffen. Gleiches gilt für Plenarprotokolle und Drucksachen der Landtage, die über das gemeinsame Informationssystem „Parlamentsspiegel.de“ recherchierbar und abrufbar sind. In allen anderen Fällen bilden die Webseiten der Institutionen die Recherchebasis. Ein Überblick über die Zugriffsoptionen kann der Excel-Tabelle „Dokumentenanalyse_Suchtrefferliste“ (Sheet 2: „Institutionenübersicht“) entnommen werden.

(1) Recherche

Da keine gemeinsame Datenbank für alle oben aufgeführten Institutionen existiert, wird – sofern vorhanden – auf die jeweilige Datenbank bzw. das Archiv einer jeden einzelnen Institution zurückgegriffen. Kann weder auf eine Datenbank, noch auf ein Archiv zugegriffen werden, so bildet die Webseite der Institution die Recherchebasis. Parallel zur Recherche erfolgt die Dokumentation der Suchergebnisse (s. Abschnitt 2 Dokumentation).

Suchbegriffe

Zur Gewährleistung einer systematischen Recherche wurden auf Basis der *PICOC*- und der *Pearl-Growing-Methode* zunächst die oben genannte Forschungsfrage abgeleitet und im Anschluss Suchbegriffe festgelegt (Booth et al., 2016; Higgins et al., 2019; Petticrew & Roberts, 2006). Diese wurden anschließend in zwei Komponenten unterteilt.

Komponente I	**Komponente II**
„Künstlich* Intelligen*“	Bildung*
KI	Hochschul*
Algorithm*	Schul*
„Learning Analtyics“	Lehr*
	Lern*

Mithilfe von Trunkierungen und booleschen Operatoren können die Suchbegriffe in zwei kombinatorische Einheiten und zu einem Suchstring verbunden werden, der die Basis für die Expertensuche im DIP bildet. Diese Suchbegriffe bilden die Basis der Recherche, können jedoch nicht für einfache Webseitensuchen in Kombination gesetzt werden. Daher werden nachfolgend die Recherchestrategien für jede einzelne Datenbank/Webseite aufgeführt.

Dokumentations- und Informationssystem für Parlamentsmaterialien (DIP)
Im Falle des DIP wurden neben den Vorgaben durch die Suchbegriffe und den Untersu-chungszeitraum zudem zwei weitere, ergänzende Einschränkungen vorgenommen. Zunächst wird die Dokumentenauswahl auf ‚Vorgänge' beschränkt. Vorgänge bilden einen von drei Informationstypen, neben Dokumenten und Aktivitäten, im DIP. Diese Einschränkung führt zu einer erheblichen Reduktion der Suchtreffer (9.813 vs. 1.758 – Stand: 15.08.2022) und ermöglicht einen besseren Überblick über die einzelnen Dokumente, indem eine, an späterer Stelle erfolgende, Entscheidung über den Ein- oder Ausschluss des Dokuments auf Basis des Vorgangstitels und -abstracts gefällt werden kann. Erweist sich ein gesamter Vorgang als irrelevant zur Beantwortung der Forschungsfrage (s. Schritt 3: Dokumentenauswahl), so kann daraus geschlossen werden, dass auch keines der darunter beratenen Dokumente (Anträge, Plenarprotokolle, etc.) als relevant erachtet werden kann. Wird ein Vorgang als relevant kategorisiert, so kann direkt auf alle dem Vorgang zugeordneten Dokumente zugegriffen werden.

Darüber hinaus wird die Auswahl auf die drei nachfolgend angeführten Sachgebiete beschränkt, um eine thematisch sinnvolle Zuordnung des Hauptthemas des Vorgangs zu gewährleisten: (1) Wissenschaft, Forschung und Technologie, (2) Medien, Kommunikations- und Informationstechnik und (3) Bildung und Erziehung.

Für die Expertensuche des DIP ergibt sich demnach der folgende Suchstring:

```
t:v AND ("künstlich* Intelligen*" OR KI OR Algorithm* OR "Learning Analytics") AND (Bildung* OR Hochschul* OR Schul* OR Lehr* OR Lern*) AND (sg:"Wissenschaft, Forschung und Technologie" OR sg:"Medien, Kommunikation und Informationstechnik" OR sg:"Bildung und Erziehung")
```

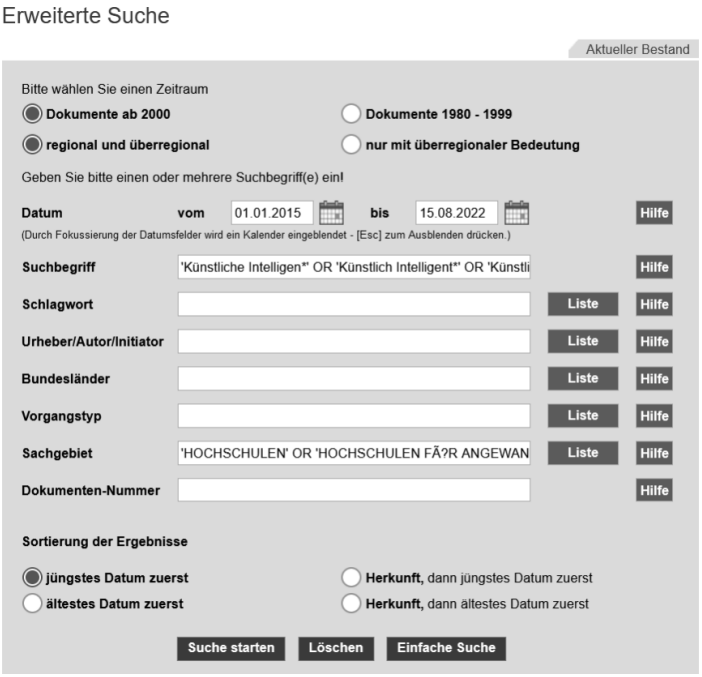
Dieser kann für die Expertensuche im DIP kopiert und ausgeführt werden. Zusätzlich muss jedoch der oben genannte Untersuchungszeitraum als weitere Filteroption erfasst werden. Zur Dokumentation (s. Schritt 2: Dokumentation) kann eine vollständige Suchtrefferliste als Word- oder PDF-Dokument heruntergeladen werden.

Parlamentsspiegel.de
Auch der Parlamentsspiegel bietet die Option einer erweiterten Suche, der Suchstring muss jedoch in diesem Fall dem System entsprechend angepasst werden. Die dort voreingestellten Filterfunktionen bleiben gesetzt. Als Datum wird der Untersuchungszeitraum angegeben. Unter Suchbegriff wird der folgende Suchstring eingetragen:

'Künstliche Intelligen*' OR 'Künstlich Intelligent*' OR 'Künstlicher Intelligenz' OR KI OR Algorithm* OR 'Learning Analytics'

Und unter Sachgebiet wird die nachfolgende Einschränkung festgelegt:

'HOCHSCHULEN' OR 'HOCHSCHULEN FÃ?R ANGEWANDTE WISSENSCHAFTEN' OR 'HOCHSCHULEN FÜR ANGEWANDTE WISSENSCHAFTEN' OR 'HOCHSCHULWESEN' OR 'BILDUNG' OR 'SCHULEN' OR 'TECHNOLOGIE'



Büro für Technikfolgenabschätzung (TAB)
Die Dokumentrecherche des TAB erfolgt zunächst über die Publikationssuche auf der Webseite der Institution. Da jedoch keine spezifischen Suchanfragen möglich sind, werden alle Suchbegriffe der ersten Komponente in das Suchfeld eingegeben und alle Suchtreffer, die im genannten Untersuchungszeitraum publiziert wurden, aufgegriffen. Anschließend erfolgt die analoge Suche im Blog des TABs.

Gemeinsame Wissenschaftskonferenz (GWK)
Die Webseite der GWK bietet die Möglichkeit, auf alle von ihr publizierten Materialien zuzugreifen. Da jedoch keine Stichwortsuche möglich ist, erfolgt die Dokumentenrecherche über den Aufruf der im Untersuchungszeitraum publizierten Dokumente und eine Volltextsuche (bei Windows über die Tastenkombination [Strg] + [F]) aller Elemente der ersten Komponente der Suchbegriffe. Kommt mindestens einer der Begriffe vor, so wird das Dokument zur weiteren Analyse einbezogen und dokumentiert (s. Schritt 2). Neben dem hier beschriebenen Einbezug der Berichte, erfolgt in einem zweiten Schritt eine Webseitensuche, um beispielsweise Dokumententypen wie Pressemitteilungen ebenfalls aufgreifen zu können. Auch hier wird die erste Komponente der Suchbegriffe verwendet.

Ständige Konferenz der Kultusminister (KMK)
Die Recherche der Dokumente der KMK erfolgt dreigeteilt. Zunächst über den Reiter „Dokumentation/Statistik“ > „Beschlüsse und Veröffentlichungen“ > „Bildung in der digitalen Welt“ und anschließend über „Dokumentation/Statistik“ > „Broschüren“ > „Bildung in der digitalen Welt“. Hier werden alle Dokumente berücksichtigt, die im Untersuchungszeitraum liegen. Im dritten Schritt werden zudem alle Pressemitteilungen des Pressearchivs einbezogen, die durch die erste Komponente der Suchbegriffe ausfindig gemacht werden können.

Expertenkommission für Forschung und Innovation (EFI)
Die Recherche der EFI-Dokumente erfolgt ebenfalls mehrteilig. Zuerst werden alle im Untersuchungszeitraum publizierten Gutachten berücksichtigt. Anschließend werden die Policy Briefs der EFI dokumentiert. Während die Gutachten zunächst vollständig mitaufgenommen werden, sofern sie im Untersuchungszeitraum veröffentlicht wurden, entscheidet bei den Policy-Briefs die thematische Fokussierung auf den Bildungssektor über die Berücksichtigung. Als Leitlinie werden hierfür die Begriffe der zweiten Komponente des Suchbegriffs herangezogen.

Wissenschaftsrat (WR)
Beim WR wird die Publikationssuche verwendet. Auch hier werden als Suchbegriffe alle Begriffe der ersten Komponente verwendet und alle darunterfallenden Dokumente, die im Untersuchungszeitraum liegen, aufgegriffen.

Hochschulrektorenkonferenz (HRK)
Für die Dokumentenrecherche der HRK wird zunächst die Publikationssuche verwendet und alle Suchbegriffe der ersten Komponente eingegeben. Da diese keine Ergebnisse aufweisen kann (Stand: 12.08.2022), wird zusätzlich die Webseitensuche, ebenfalls mit allen Begriffen der ersten Komponente, genutzt. Dabei werden alle Dokumente herausgegriffen, die unter den Typ „Pressemitteilung“ oder „Beschlüsse“ fallen (Filteroption am linken Rand).

Hochschulforum Digitalisierung (HFD)
Hier erfolgt die Recherche erneut zweiteilig. Zunächst werden in der Publikationssuche alle Begriffe der ersten Komponente aufgegriffen. Anschließend werden alle Beiträge der Themendossiers „Künstliche Intelligenz“ und „Learning Analytics“ berücksichtigt. Dies beinhaltet explizit auch Blogbeiträge oder andere Publikationstypen.

Gewerkschaft Erziehung und Wissenschaft (GEW)
Bei der GEW wird eine einfache Webseitensuche mit den Begriffen der ersten Komponente durchgeführt. Die Webseitensuche bietet im Vergleich zur Publikationssuche den Vorteil, dass neben als Publikationen deklarierten Dokumenten auch beispielsweise Nachrichten oder Pressemitteilungen berücksichtigt werden. Suchtreffer, die als „Veranstaltung“ gekennzeichnet sind, werden außen vor gelassen.

Hochschullehrerbund (HLB)
Die Webseitensuche wird ebenfalls für den HBL verwendet. Dokumentiert werden alle Suchergebnisse der Recherche der ersten Komponente im Untersuchungszeitraum.

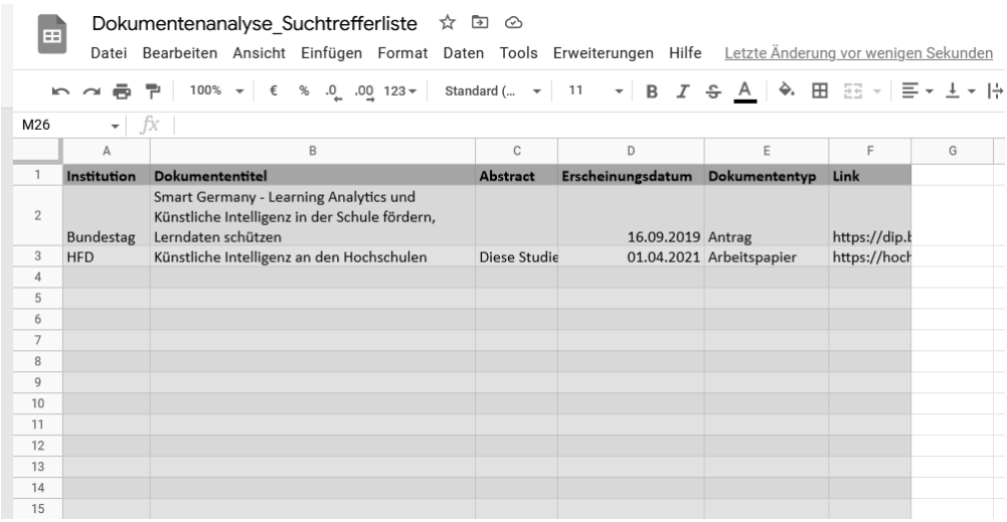
Deutscher Lehrerverband (DL)
Auch beim DL wird eine Webseitensuche mit den Suchbegriffen der ersten Suchkomponente durchgeführt.

(2) Dokumentation

Parallel zur Recherche erfolgt die Dokumentation aller Suchtreffer. Dafür wird in der auf Google Drive abgelegten Excel-Tabelle „Dokumentenanalyse_Suchtrefferliste“ (Sheet 1: „Suchtrefferliste“) eine vollständige Liste aller durch die oben beschriebenen Verfahren gefundenen Dokumente erstellt. Festgehalten werden dort die folgenden formalen Kategorien: Institution, Dokumententitel, Abstract (wenn vorhanden), Erscheinungsdatum, Dokumententyp (z. B. Antrag, Pressemitteilung, Plenarprotokoll, Blogbeitrag) und Link. Das Erscheinungsdatum wird in der Form TT.MM.JJJJ angegeben. Ist kein konkretes Publikationsdatum ersichtlich, sondern nur ein Monat und/oder ein Jahr, so wird jeweils der 01. des Monats/des Jahres angegeben.

Da aufgrund der Vielzahl an Datenquellen kein automatisierter Download möglich ist, werden aus forschungsökonomischen Gründen an dieser Stelle noch nicht alle Dokumente heruntergeladen und archiviert, sondern lediglich der Zugangslink dokumentiert. Der Download erfolgt somit erst nach der finalen Auswahl der Dokumente für die anschließende Analyse (siehe Abschnitt 3).

Für die Suchergebnisse des DIP gelten ein paar Besonderheiten: Da sich durch die Einschränkung der Parlamentsmaterialienrecherche auf Vorgänge kein einzelnes Erscheinungsdatum benennen lässt, wird in diesem Fall das Erscheinungsdatum der ersten Aktivität des Vorgangs angegeben. Unter „Dokumententitel“ wird der Name des Vorgangs angegeben. Zudem wird unter Dokumententyp die Art des Vorgangs dokumentiert, also z. B. Antrag, schriftliche Frage, kleine Anfrage, Gesetzgebung.



	A	B	C	D	E	F	G
1	**Institution**	**Dokumententitel**	**Abstract**	**Erscheinungsdatum**	**Dokumententyp**	**Link**	
2	Bundestag	Smart Germany - Learning Analytics und Künstliche Intelligenz in der Schule fördern, Lerndaten schützen		16.09.2019	Antrag	https://dip.b	
3	HFD	Künstliche Intelligenz an den Hochschulen	Diese Studie	01.04.2021	Arbeitspapier	https://hoch	
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							

(3) Dokumentenauswahl

Basierend auf der Gesamtzahl der dokumentierten Suchtreffer erfolgt die Auswahl der Dokumente für die Analyse. Der Schritt der Dokumentenauswahl verfolgt das Ziel, die Dokumente zu identifizieren, die einen Beitrag zur Beantwortung der Forschungsfrage liefern. Die Reduktion der Dokumente erfolgt dabei mehrstufig.

Um die Relevanz eines Beitrags für die anschließende Analyse beurteilen zu können, werden die Dokumente von insgesamt drei Gutachtern bewertet. Dafür wird die Suchtrefferliste zunächst in eine zufällige Reihenfolge gebracht. Anschließend werden zum Zwecke einer ersten Intercoderreliabilitätsprüfung die ersten 50 Dokumente der Suchtrefferliste von allen drei Gutachtern bewertet.[2] So sollen die Codier-Ergebnisse statistisch abgesichert sowie die Qualität der Auswertung erhöht werden. Die Bewertung der restlichen Dokumente erfolgt in zufälliger Reihenfolge, aufgeteilt zwischen den drei Gutachtern.

3.1 Zugriff

Für die Dokumentenauswahl wird zuerst der Zugriff auf das Dokument überprüft. Ist der Zugriff auf ein Dokument nicht möglich, so gilt dies als Ausschlusskriterium. Für die Variable „*Zugriff*“ werden demnach die folgenden Codes vergeben:

Zugriff	[0]	Kein Zugriff möglich
	[1]	Zugriff möglich

Die Variable Zugriff gilt darüber hinaus als *Filtervariable*: Ist der Zugriff nicht möglich, so wird die Codierung an dieser Stelle beendet.

[2] Insgesamt wurden drei Reliabilitätstests mit einer Gesamtzahl von 131 Dokumenten durchgeführt. Der abschließende Reliabilitätstest (n = 31) ergab eine durchschnittliche Übereinstimmung von 0,96. nach Holsti.

3.2 Relevanz

Rel_a

[Hinweis: Da zwischen allen drei Gutachtern nach Abschluss des zweiten Reliabilitätstests Einigkeit herrschte, dass die Codierung von Rel_a keinen gewinnbringenden Beitrag für die Beurteilung eines Dokuments in Hinblick auf den Mehrwert zur Beantwortung der Forschungsfrage leistet, wurde diese Variable für die weiteren Codierungen ausgeschlossen.]

Für die Beurteilung der Relevanz werden zunächst der in der Suchtrefferliste dokumentierte Titel und das Abstract (falls vorhanden) des Dokuments herangezogen – sind Schlagworte vorhanden (i. d. R. bei Beiträgen des DIPs der Fall), so werden auch diese berücksichtigt. Auf dieser Basis wird beurteilt, ob das vorliegende Dokument die Thematik von KI im Bildungssektor mindestens im erweiterten Sinne behandelt. Dies schließt beispielsweise Dokumente mit ein, die Digitalisierungs- oder Technologie-Themen in der öffentlichen Verwaltung, in der (inner- und außeruniversitären) Forschung oder anderen Bildungsbereichen (wie z. B. frühkindliche oder außerschulische Bildung) mit ein. Durch die zuvor festgelegten Suchkriterien und die Auswahl der Institutionen sollte dies in der Regel gewährleistet sein. Daher wird der Code [1] vergeben, wenn der Titel, das Abstract oder die Schlagworte des Dokuments einen erkennbaren Bezug zur ersten oder zweiten Komponente der Suchbegriffe aufweisen. Die Suchbegriffe müssen nicht explizit vorkommen, ein thematischer Zusammenhang sollte jedoch bestehen. Da die spätere Inhaltsanalyse auf Aussagenebene erfolgt, kann die Variable Rel_a „großzügig" behandelt werden: In diesem Schritt geht es demnach erst einmal nur darum, gegebenenfalls irrtümlich aufgenommene Dokumente auszusortieren, die keinerlei erkennbaren Bezug zu digitalen Technologien im Bildungskontext aufweisen. So kann beispielsweise bereits ein Verweis auf Weiterbildungskurse, Ausbildung, Forschung oder Digitalisierung als thematisch einschlägig gewertet werden.

Rel_a	[0]	Thematisch nicht einschlägig
	[1]	Thematisch einschlägig

Besonderheit: Wurde zuvor unter Zugriff der Code [0] vergeben, d. h., der Zugriff auf das Dokument ist nicht möglich, und bei der Bewertung der Variable Rel_a jedoch der Code [1] vergeben, so wird der Titel des Dokuments in der Tabelle grün markiert, sodass das Dokument nachträglich manuell gesucht werden kann.

Rel_b

Die weitere Relevanzbeurteilung erfolgt auf Basis einer Volltextsuche. Konkret bedeutet das, dass das zu beurteilende Dokument zunächst aufgerufen und anschließend eine Volltextsuche der ersten Komponente der Suchbegriffe sowie der Begriffe „blended learning" und „e(-)Learning" durchgeführt wird. Hinzu kommen außerdem „maschinelles Lernen", „neuronale Netze" sowie „automatisierte Systeme". Zur schnelleren Durchsicht empfiehlt es sich, die Suchbegriffe in der folgenden Form in der Volltextsuche zu verwenden:

künstlich – KI – Algo – learn – maschin – neuron – automat

Das Ergebnis dieser Volltextsuche bildet die Grundlage für die Bewertung der nachfolgenden Variable, die das Vorkommen mindestens eines der Begriffe prüft:

Rel_b	[0]	Keiner der Begriffe kommt vor
	[1]	Mindestens ein Begriff kommt vor

Besonderheit: Da sich die Recherche im DIP auf gesamte Vorgänge bezieht, werden dort häufig mehrere Dokumente angegeben. Die Volltextsuche wird hier für jedes einzelne Dokument durchgeführt. Handelt es sich jedoch um ein Plenarprotokoll, so werden nur die Suchergebnisse in der im Vorgang angegebenen Seitenspanne des Plenarprotokolls berücksichtigt. Gleiches gilt für Plenarprotokolle aus Landtagen. Hier wird sich ebenfalls nur auf den Abschnitt des

Dokuments bezogen, der sich auf den in der Suchtrefferliste festgeschriebenen Dokumententitel bezieht.

Rel_c

Da die nachfolgend angestrebte Dokumentenanalyse auf Aussagenebene durchgeführt wird, ist es nicht zwingend notwendig, dass KI im Bildungssektor das Hauptthema des Dokuments ist. Es ist stattdessen vollkommen ausreichend, wenn einzelne Sätze des Dokuments einen Beitrag zur Beantwortung der Forschungsfrage leisten. Daher werden im nächsten Schritt die Absätze als Bewertungsgrundlage der Relevanz herangezogen, die durch die vorherige Volltextsuche identifiziert wurden und somit mindestens einen der Suchbegriffe enthalten. Die Absätze werden gescannt und in Hinblick auf einen Beitrag zur Beantwortung der oben aufgeführten Forschungsfrage hin überprüft. Einen Hinweis auf einen Beitrag zur Beantwortung der Forschungsfrage bietet beispielsweise das **Vorkommen einer beliebigen Form der Erklärung, die Nennung eines Arguments (für oder gegen) eine Anwendung**, einen Einsatz oder beispielsweise die Forschung und Auseinandersetzung an und mit KI im Bildungssektor oder eine **andere Form der Rechtfertigung**. Der Bezug zu KI wird dabei in einem sehr weiten Sinne verstanden, sodass der Code [1] den Fokus auf die Form der Äußerung, i. S. einer Erklärung, eines Arguments, einer Rechtfertigung oder einer anderen Form der Begründung einer Forderung oder eines Handelns, legt und nicht auf das Objekt der Begründung (z. B. „die KI“, der Chatbot, die Forderung von KI-Studiengängen). Demnach wird der Code [1] auch dann vergeben, wenn zum Beispiel die Forderung von Forschungsförderung im Zusammenhang mit KI im Bildungssektor begründet wird.

Rel_c	[0]	Kein Bezug zur Forschungsfrage
	[1]	Bezug zur Forschungsfrage

Rel_d

Im letzten Schritt wird schließlich codiert, ob **konkrete** (z. B. Performance-Prediction-Systeme, Learning Analytics, Dropout-Detection) **oder diffuse KI-Anwendungen** (i. S. einer allgemeinen Nennung von KI-Anwendungen anstelle der Nennung konkreter Anwendungen), die im Bildungssektor zum Einsatz kommen, thematisiert werden oder ob über andere Erscheinungsformen von KI berichtet wird (z. B. KI allgemein, KI-Forschung). Dabei ist es unerheblich, in welchem Bereich des Bildungssektors die KI-Anwendungen zum Einsatz kommen (sollen) – wie beispielsweise in der Lehre, in der (öffentlichen) Verwaltung oder in der Forschung – oder ob sie bereits eingesetzt werden, in der Vergangenheit wurden oder über einen zukünftigen Einsatz diskutiert wird.

Rel_d	[0]	Kein Bezug zu KI-Anwendung(en)
	[1]	Bezug KI-Anwendung(en)

Wenngleich die spätere Analyse auf Aussagenebene durchgeführt wird, werden die hier zu vergebenen Codes dennoch auf Dokumentenebene vergeben, da dieser Schritt zunächst nur der finalen Dokumentenauswahl dient und keinesfalls die tiefgreifendere Analyse der gegebenen Rechtfertigungen ersetzt.

Relevanzindex

Zur finalen Auswahl der Dokumente wird ein Summenscore über die drei Relevanzkriterien (Rel_b, Rel_c, Rel_d) gebildet. Vorab werden jedoch, wie bereits beschrieben, alle Dokumente ausgeschlossen, auf die kein Zugriff möglich ist. Auf der Basis des Summenscores kann ein Ranking der Bedeutsamkeit der Dokumente zur Beantwortung der Forschungsfrage erstellt werden. Alle Dokumente, die einen Wert von drei erhalten, werden für die anschließende Inhaltsanalyse berücksichtigt. Im Sinne einer theoretischen Sättigung wird nach Abschluss der Codierung dieser Dokumente entschieden, ob die Dokumente, die einen Wert von zwei erreicht haben, ebenfalls in die Analyse aufgenommen werden. Alle Dokumente, die einen Wert von eins erreichen, werden für die Analyse nicht berücksichtigt.

(4) Download

Der Download und die Archivierung erfolgen aus forschungsökonomischen Gründen nur für die Dokumente, die im Relevanzindex einen Wert von drei erreichen. Ein erfolgter Download wird in der Spalte „Download" der Excel-Tabelle „Dokumentenanalyse_Suchtrefferliste" mit einer [1] dokumentiert. Der Download erfolgt in der folgenden Form:

Downloadformat: Die Dokumente stehen in der Regel als PDF zur Verfügung und können auch in diesem Format heruntergeladen werden. Steht kein PDF zum Download zur Verfügung, bspw. weil es sich um einen Webseitentext handelt, so wird das PDF selbst erstellt. Dies kann beispielsweise über die Druckfunktion (bei Windows über die Tastenkombination [Strg] + [P], Auswahl: Ziel = Adobe PDF) erfolgen.

Speichertitel: Die Dokumente werden alle nach dem nachfolgenden Schema benannt:

Kürzel der Institution (s. o.) _ Erscheinungsdatum (TT.MM.JJJJ) _fortlaufende Nummer _ ggf. Kleinbuchstabe (fortlaufend)

Das Erscheinungsdatum sowie die fortlaufende Nummer werden der Excel-Tabelle „Dokumentenanalyse_Suchtrefferliste" entnommen. Da es sich bei den Suchtreffern des DIPs jedoch um Vorgänge handelt, unter die jeweils mehrere Dokumente fallen können, wird dort hinter der vorlaufenden Nummer zusätzlich ein Kleinbuchstabe ergänzt.

Beispiele: BT_26.06.2018_113_a.pdf; BT_26.06.2018_113_b.pdf; EFI_12.01.2015_1078

(5) Archivierung

Zur Archivierung der gedownloadeten PDF-Dokumente wird eine ZIP-Datei angelegt und per E-Mail an birte.keller@hhu.de geschickt. Anschließend werden diese ZIP-Dateien auf dem Laufwerk des Lehrstuhls abgelegt.

Literaturverzeichnis

Booth, A., Sutton, A. & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review* (Second edition). SAGE.

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B. & Ullmann, Thomas and Vuorikari, Riina. (2016). *Research Evidence on the Use of Learning Analytics: Implications for Education Policy.* Joint Research Centre Science for Policy Reports. https://doi.org/10.2791/955210

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J. & Welch, V. A. (Hrsg.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (Second edition). John Wiley & Sons.

Ifenthaler, D. & Schumacher, C. (2016). Learning Analytics im Hochschulkontext. *WiSt - Wirtschaftswissenschaftliches Studium*, *45*(4), 176–181. https://doi.org/10.15358/0340-1650-2016-4-176

Petticrew, M. & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell Publishing.

Willems, C. (2021). *Künstliche Intelligenz im politischen Diskurs*. Tectum Verlag.

Reliabilitätstest-Nr.	Variable	Bezeichnung	Koeffizient	Coderer:innenpaar A/B	A/C	B/C	Durchschnittliche Übereinstimmung
Reliabilitätstest I (n = 50)	Zugriff	Dokumentzugriff möglich	Holsti r_H	0,980	0,960	0,980	0,973
			Cohen‘s Kappa r_K	0,847	0,728	0,847	0,807
	rel_a	Thematisch einschlägig	Holsti r_H	0,640	0,800	0,520	0,653
			Cohen‘s Kappa r_K	0,280	0,600	0,052	0,311
	rel_b	Vorkommen der Suchbegriffe	Holsti r_H	0,940	0,920	0,900	0,920
			Cohen‘s Kappa r_K	0,855	0,822	0,769	0,815
	rel_c	Bezug zur Forschungsfrage	Holsti r_H	0,820	0,660	0,760	0,747
			Cohen‘s Kappa r_K	0,593	0,101	0,339	0,345
	rel_d	KI-Anwendung im Bildungskontext	Holsti r_H	0,900	0,840	0,900	0,880
			Cohen‘s Kappa r_K	0,558	0,115	0,494	0,389
Durchschnittliche Übereinstimmung			**Holsti r_H**	**0,856**	**0,836**	**0,812**	**0,835**
			Cohen‘s Kappa r_K	0,627	0,473	0,500	0,533
Reliabilitätstest (n = 50)	Zugriff	Dokumentzugriff möglich	Holsti r_H	0,960			
	rel_a	Thematisch einschlägig	Holsti r_H	0,600			
	rel_b	Vorkommen der Suchbegriffe	Holsti r_H	0,940			
	rel_c	Bezug zur Forschungsfrage	Holsti r_H	0,700			
	rel_d	KI-Anwendung im Bildungskontext	Holsti r_H	0,800			
Durchschnittliche Übereinstimmung			**Holsti r_H**	**0,800**			
Reliabilitätstest III (n = 31)	Zugriff	Dokumentzugriff möglich	Holsti r_H	0,935	0,935	1,000	0,957
			Cohen‘s Kappa r_K	0,475	0,457	1,000	0,650
	rel_a	Thematisch einschlägig		-	-	-	-
	rel_b	Vorkommen der Suchbegriffe	Holsti r_H	0,968	0,968	1,000	0,979
			Cohen‘s Kappa r_K	0,935	0,935	1,000	0,957
	rel_c	Bezug zur Forschungsfrage	Holsti r_H	0,903	0,903	0,903	0,903
			Cohen‘s Kappa r_K	0,739	0,687	0,739	0,722
	rel_d	KI-Anwendung im Bildungskontext	Holsti r_H	1,000	1,000	1,000	1,000
			Cohen‘s Kappa r_K	1,000	1,000	1,000	1,000
Durchschnittliche Übereinstimmung			**Holsti r_H**	**0,952**	**0,952**	**0,976**	**0,960**
			Cohen‘s Kappa r_K	0,787	0,770	0,935	0,832

A4. Codierleitfaden

Codierleitfaden der inhaltlich strukturierenden qualitativen Inhaltsanalyse

im Rahmen der Dokumentenanalyse des Dissertationsprojekts von Birte Keller

Über die Legitimation der Einführung künstlich intelligenter Systeme

Hinweis: Dieses Dokument und insbesondere das Kategoriensystem sowie theorie- und untersuchungsgegenstandserläuternde Textpassagen sind als Bestandteil der Dissertation streng vertraulich zu behandeln und nicht ungefragt an Dritte weiterzugeben.

Glossar & Abkürzungsverzeichnis

Begriff	Definition
Cité	Unterschiedliche Gerechtigkeitsvorstellungen innerhalb einer Gesellschaft, die als moralische Grundsätze und somit als rechtfertigende Gründe für das eigene Handeln herangezogen werden; Sind nicht selbst als Rechtfertigungen zu verstehen, sondern vielmehr als Äquivalenzprinzipien, die eine (moralische) Grundlage einer Rechtfertigung bilden; auch als Rechtfertigungsordnung, Polis oder Gemeinwesen bezeichnet.
Codes	Markierungen von Textstellen, die in MAXQDA am Text vorgenommen werden und einzelnen Kategorien zugeordnet werden
Kategorie	Wichtigstes analytisches Instrument der QIA; dient der Bedeutungszuschreibung, Klassifizierung, Reduzierung und Abstrahierung
QIA	Qualitative Inhaltsanalyse
Rechtfertigung	Rechtfertigungen begründen das eigene (Nicht-)Handeln, einen persönlichen Standpunkt, eine Forderung oder Kritik; sie können positive Argumente, Behauptungen oder Stellungnahmen sein, aber auch kritische ‚Anprangerungen' gegnerischer Ansichten (Thévenot et al., 2000, S. 237).
Segment	Textabschnitt, der eine in sich abgeschlossene Sinneinheit innerhalb eines Dokuments umfasst und als Grundlage, d. h. Analyseeinheit der QIA gilt.
Sinneinheit	Eine Sinneinheit bezeichnet den kleinsten inhaltlich abgeschlossenen Textabschnitt, der eine eigenständige Aussage oder Argumentation enthält und daher als Segment definiert wird, d. h. als Analyseeinheit der Codierung dient. Sie umfasst typischerweise einen vollständigen Gedanken, der unabhängig von seiner formalen Länge in sich kohärent ist.
Suchbegriffe	Begriffe, die sich mit dem Gegenstand KI befassen und im Rahmen des systematischen Vorgehens der Dokumentenauswahl definiert wurden. Diese umfassen u. a. die folgenden Begriffe inkl. ihrer Abwandlungen: künstliche Intelligenz, KI, Algorithmen, (Machine, Deep) Learning (Analytics), maschinell(es Lernen), neuronale Netze, automatisierte Systeme

Forschungsfragen

Übergeordnete Forschungsfragen der Dokumentenanalyse:

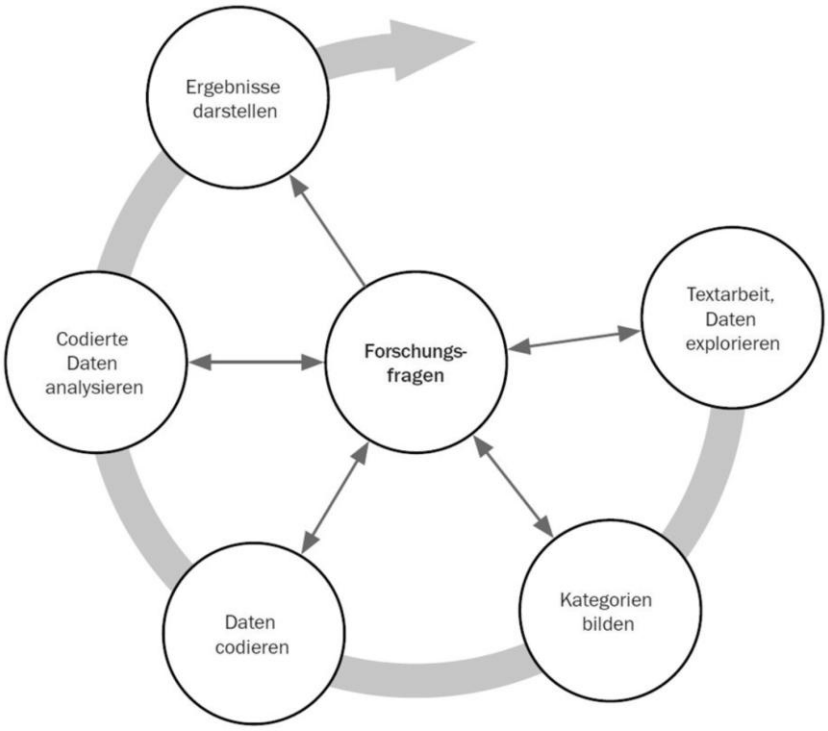
1. Welche Rechtfertigungen äußern bildungspolitische Akteur:innen für oder gegen den Einsatz von künstlich intelligenten Systemen im deutschen Bildungssystem?
2. Welche Rechtfertigungsordnung dominiert die Debatte des Einsatzes von künstlich intelligenten Systemen im deutschen Bildungssystem?

Untergeordnete Forschungsfragen der Dokumentenanalyse:

- Welche Rechtfertigungen werden für den Einsatz von KI-Systemen im deutschen Bildungssystem geäußert?
- Welche Rechtfertigungen werden gegen den Einsatz von KI-Systemen im deutschen Bildungssystem geäußert?
- Welche Akteur:innen äußern diese Rechtfertigungen für oder gegen den Einsatz von KI-Systemen im deutschen Bildungssektor?
- Welche Art(en) von KI-Systemen wird/werden gerechtfertigt?
- Werden KI-Systeme, die in unterschiedlichen Bereichen (z. B. Lernen vs. Verwaltung) oder auf unterschiedlichen Stufen (z. B. Primär vs. Tertiär) des Bildungssystems zum Einsatz kommen, unterschiedlich gerechtfertigt?

Vorgehen der qualitativen Inhaltsanalyse (QIA)

Zur Beantwortung der oben genannten Forschungsfragen wird ein inhaltlich-strukturierender Ansatz der QIA gewählt (vgl. Kuckartz & Rädiker, 2022; Mayring, 2022). Diese orientiert sich an dem folgenden schematischen Ablauf:



Quelle: Kuckartz & Rädiker, 2022, S. 106

Allgemeine Hinweise zur Arbeitsorganisation

Zur Durchführung der QIA wird ein über die **TeamCloud** organisiertes, gemeinsames MAXQDA-Projekt verwendet. Zur Verwendung von MAXQDA ist eine bestehende VPN-Verbindung zum Server des IKMs notwendig. Um Änderungen des gemeinsamen Projekts herunter- oder hochzuladen, muss man in der TeamCloud angemeldet sein, die lokale Bearbeitung des Teamprojekts kann aber auch offline erfolgen (*Hinweise zur Installation s. \GitHub\dissertation-ki-legitimation\Dokumentenanalyse\MAXQDA*). Vor Beginn und nach Abschluss einer jeden Arbeitssession muss der aktuelle Arbeitsstand herunter- bzw. hochgeladen werden.

Besonderheiten, vorgenommene Änderungen und Fortschritte sowie Probleme und Fragen werden im Projekt-Logbuch (Reiter „TeamCloud“) schriftlich festgehalten. Das **Projekt-Logbuch** wird von jedem Codierer/jeder Codiererin vor Start einer Analysesession konsultiert. Rückfragen und Diskussionen werden über Basecamp kommuniziert. Das Logbuch kann als Ergebnisprotokoll getroffener Entscheidungen verstanden werden. Besonderheiten einzelner Dokumente, Segmente oder Codes werden als „Memos“ direkt an der jeweiligen Stelle vermerkt.

Die Daten, d. h. die in einem vorherigen Schritt systematisch ausgewählten Dokumente, sind bereits in dem Projekt enthalten und nach Institutionen gruppiert. Eine zufällige Zuordnung der Dokumente zu den drei Codierer:innen ist erfolgt und die jeweils einer/einem Codierer:in zugeordneten Dokumente wurden zu einem „Set“ zusammengefügt.

Datenexploration

Die Datenexploration dient dazu, die später im Detail zu codierenden Segmente (i. S. v. Textelementen und Sinneinheiten) in den Dokumenten zu identifizieren. Zu diesem Zweck wird ein teilautomatisiertes Vorgehen gewählt und eine erweiterte Textsuche und Autocodierung durchgeführt (*weiterführende Hinweise im Manual: https://www.maxqda.com/de/hilfe-mx22/lexikalische-suche-2/die-erweiterte-lexikalische-suche*). Diese ermöglicht sowohl den Ein- als auch den Ausschluss einzelner Wörter und Wortgruppen für die Textsuche. Als Orientierung gelten die zur Dokumentenauswahl genutzten Suchbegriffe:

Künstlich* intelligen* – KI – K. I. – Algorithm* – learn* – maschin* – neuron* – automat*

Das Ergebnis dieser Suche identifiziert alle Textstellen in den Dokumenten, die diese Suchbegriffe enthalten. Im nächsten Schritt wurden fehlerhafte Begriffe ausgeschlossen, d. h. eben jene Begriffe, die keine Relevanz zur Beantwortung der Forschungsfragen aufweisen. Dafür wurden die Suchbegriffe gesichtet und fehlerhafte Begriffe für die nachfolgende automatische Codierung als ‚zu ignorierend‘ markiert. Dazu zählen z. B.:

Maschinenbau* – Maschinenverordnung – Maschinenwartung – Neuronen – NeuroNex

Anschließend wurden alle Sätze, die mindestens einen der aufgeführten Suchbegriffe enthielten, automatisch der Kategorie „A. 1 Suchbegriffe“ zugeführt.[1] Diese wurden zudem in die

[1] MAXQDA > Textsuche & Autocodieren: Suchbegriffe: künstl* intelligen* KI KI- K.I. K. I. learn* -learn* elearn* maschin* neuron* automat* algorithm* Avatar | Einstellungen: Innerhalb von: Dokumenten; Nur ganze Wörter; Finde alle Wortformen (Lemma-Liste): German | Ausschluss von irrelevanten Begriffen: algorithmic_accountability_on_the_investigation_of_black_box, algorithmusverbesserterwerbschancenvonfluechtlingen, AlgorithmWatch_AutoHR_Gutachten_Arbeitsrecht_Datenschutz_Wedde_2659, AlgorithmWatch_AutoHR_Positionspapier_2020, AlgorithmWatch_AutoHR_Study_Ethics_Loi_2020, Automatenaufstellers, Automatenhersteller, Automatenspiel, Automating_Society_Report_2019, Automatisch (wenn thematisch irrelevant), Automatismus, ki- (HTML-Links), Learnweb, Maschinenbau*, Maschinenbediener, Maschinenpark, Maschinenrings, Maschinenschlosser, Maschinentyp, Maschinenwartung, NeuroNex

Kategorie „A.2 Segmente“ kopiert und dienen als Analysegrundlage der weiteren QIA. So kann nun jede:r Codierer:in kann auf diese Textstellen zugreifen und diese als Orientierungspunkt zur Definition der Analyseeinheit nutzen. Das bedeutet jedoch nicht, dass nicht auch andere Textstellen der Dokumente als Analyseeinheiten definiert werden dürfen. Für die weitere initiierende Textarbeit werden die automatisch codierten Textstellen nun einzeln bewertet und ausgeweitet, sodass die Codierung eine komplette Sinneinheit umfasst. Zur Bewertung und Festlegung einer Sinneinheit werden die Forschungsfragen in Erinnerung gerufen. Sie ist zu verstehen als kleinster inhaltlich abgeschlossener Textabschnitt, der eine eigenständige Aussage oder Argumentation enthält, sodass sie einen vollständigen Gedanken umfasst, der unabhängig von seiner formalen Länge in sich kohärent ist. Sollte eine automatisiert erzeugte Codierung nicht von Relevanz für die Beantwortung der Forschungsfrage sein, so wird diese aus der Kategorie „A.2 Segmente“ gelöscht.

Wenn sinnvoll, können Memos zu einzelnen Dokumenten oder Segmenten angelegt werden, um Besonderheiten, Ideen für Kategorien oder Analysen oder erste Antworten auf Forschungsfragen zu dokumentieren.

Kategorienbildung

Die hier durchzuführende QIA verfolgt einen deduktiv-induktiven Ansatz der Kategorienbildung, d. h., es wurden bereits deduktive Kategorien aus der theoretischen Grundlage der Dissertation abgeleitet (s. u. Kategoriensystem), die jedoch im Laufe der QIA weiter induktiv ergänzt werden. Zum Verständnis der deduktiv erarbeiteten Kategorien dienen die unten in diesem Dokument angeführten Ausführungen „Zum Verständnis der Cités“, die vor Beginn der Analyse zu lesen sind. Rückfragen können jederzeit über Basecamp oder persönlich gestellt werden.

Die deduktiv erfassten Kategorien wurden im Codesystem des Projekts in MAXQDA importiert. Die Kategoriendefinition, Ankerbeispiele und Abgrenzungen finden sich in den Memos und können und sollen fortlaufend präzisiert werden.

Basiscodierung

Jede:r Codierer:in nimmt eine Basiscodierung der ihm/ihr zugeordneten Dokumente vor. Diese Basiscodierung dient der inhaltlichen Strukturierung. Dabei gilt es, Folgendes zu beachten:

- **Nicht zu kleinteilig Codieren**: Es sollen stets ganze Sinneinheiten codiert werden, die sich daraus ergeben, dass es um einen Aspekt geht. Dies kann einen Satz umfassen oder auch einen ganzen Abschnitt. Die Codiereinheit muss immer für sich alleine verständlich sein, d. h. auch ohne den sie umgebenden Kontext.
- **Eher großflächig codieren**: Nicht mehrfach hintereinander den gleichen Code für Minisätze vergeben.
- **Wiederholte Nennung beim gleichen Fall**: Üblicherweise mehrfach codieren, jedoch Kommentar, dass dies bereits mehrfach im Dokument codiert wurde.
- **Unwichtiges**: Nicht codieren.
- **Wenn für relevante Aussagen noch keine passende Hauptkategorie existiert**: Sonstiges-Kategorie verwenden oder bei besonders relevanten Aspekten sofort eine neue Kategorie ergänzen.
- **Wenn die Sonstiges-Kategorie bezüglich einzelner Aspekte zu voll wird**: Neue Kategorie ergänzen oder ggf. Unterkategorien bei „Sonstiges“ verwenden und gemeinsam diskutieren.
- **Kategoriendefinitionen**: Bei Bedarf anpassen, Abgrenzungen zu anderen Kategorien und Beispiele einfügen.

Feincodierung

Im Anschluss erfolgt die Feincodierung. Dafür werden alle Segmente einer Hauptkategorie zusammengestellt und gemeinsam diskutiert, wie mit dieser weiter vorgegangen werden soll: Sind Subkategorien sinnvoll? Wenn ja, wie können diese aussehen? Wie können die Subkategorien definiert werden?

Konsensuelles Vorgehen

- Konsensuelles Vorgehen und Austausch über Nicht-Übereinstimmungen als Gütekriterium der Intercoder-Übereinstimmung (Kuckartz & Rädiker, 2022)
- Iteratives Vorgehen:
  (1) Ein bis zwei Fälle gemeinsam codieren.
  Kategoriendefinitionen überprüfen und verbessern, Codierregeln aufstellen und ergänzen.
  (2) Ein bis zwei Fälle getrennt codieren und Codierungen vergleichen (ggf. Übereinstimmung berechnen).
  Kategoriendefinitionen und Codierregeln verbessern und erweitern.
  (3) Schritt 2 so lange wiederholen, bis eine zufriedenstellende Übereinstimmung erzielt wird.

A5. Kategoriensystem

Kategorienübersicht

Formale Kategorien/Dokumentvariablen *[bereits importiert]*

Fortlaufende Nummer	Dokumententyp
Speichertitel	Link
Haupt-/Nacherhebung	Zugriff
Codierer:in Dokumentenauswahl	Rel_b [Vorkommen der Suchbegriffe]
Rel-Test	Rel_c [Bezug zur Forschungsfrage]
Zufallszahl I [Dokumentenauswahl]	Rel_d [KI-Anwendung im Bildungskontext]
Institution	Relevanzindex
Dokumententitel	Analyserelevante Anmerkungen
Abstract	Bemerkung Codierung (Original)
Erscheinungsdatum	Zufallszahl II [Codierung]
Erscheinungsjahr	Codierer:in QIA

Inhaltliche Kategorien

Block	Kategorie
A. Analyserelevante Kategorien	A.1 Suchbegriffe
	A.2 Segmente
	A.3 Zitierfähige Aussagen
	A.4 Unklarheiten
	A.5 Neue Ideen
	A.6 Blumen am Wegesrand
B. Kontext	B.1 KI-Bezug
	B.1.1 KI-Anwendung konkret
	B.1.1.1 KI zum Lernen
	B.1.1.1.1 Blended/E-Learning
	B.1.1-2 KI in der Lehre
	B.1.1.3 KI zur Bewertung
	B.1.1.4 KI zu Forschungszwecken
	B.1.1.5 KI in der Administration
	B.1.2 KI-Anwendung diffus
	B.1.3 Wissenserwerb über KI
	B.1.4 Investition in KI
	B.1.5 Forschung über KI[1]
	B.1.6 Sonstige
	B.2 Verortung im Bildungssystem
	B.2.1 Elementarbereich
	B.2.2 Primar- & Sekundarbereich[2]
	B.2.3 Tertiärbereich
	B.2.4 Quartärbereich
	B.2.5 Unspezifizierter Bildungsbereich[1]
	~~B.3 Folgen der KI~~[3]
	~~B.3.1 Positive Folgen~~[3]
	~~B.3.2 Negative Folgen~~[3]

[1] Induktiv ergänzte Kategorie.
[2] Zu Beginn der Codierung einzelne Kategorien, im Zuge des konsensuellen Codierens jedoch zusammengefasst, da in der Regel nicht zwischen Primar- und Sekundarbereichen der Schule unterschieden wurde.
[3] Im Zuge Probecodierungen gestrichen, da keine ausreichenden Abgrenzungskriterien gegenüber der „Richtung der Rechtfertigung (C.3) formuliert werden konnten.

Block	Kategorie
C. Rechtfertigung	C.1 Cité C.1.1 Inspiration C.1.2 Haus C.1.3 Meinung C.1.4 Staatsbürger C.1.5 Industrie C.1.6 Markt C.1.7 Projekt C.1.8 Ökologie C.1.9 Solution C.1.10 Andere Form der Rechtfertigung C.1.11 Keine Rechtfertigung[1] C.1.12 Solution (implizit und explizit)[4] C.1.13 Anti-Solution[1,5] C.2 Prüfungen[1] C.2.1 Wahrheitsprüfung[1] C.2.2 Realitätsprüfung[1] C.2.3 Existenzprüfung/Konflikt[1,6] C.2.4 Kompromiss[1] C.3 Richtung der Rechtfertigung C.3.1 Pro-KI C.3.2 Abwägend Positiv C.3.3 Keine Valenz C.3.4 Abwägend Negativ C.3.5 Contra-KI C.4 Rechtfertigende:r Akteur:in C.4.1 Politische Akteur:innen C.4.1.1 Parteien C.4.1.1.1 Die Linke C.4.1.1.2 Bündnis 90/Die Grünen C.4.1.1.3 SPD C.4.1.1.4 CDU/CSU C.4.1.1.5 FDP C.4.1.1.6 AfD C.4.1.1.7 Andere Partei C.4.1.2 Regierungen[1] C.4.1.3 Parlamentarische Gremien[1] C.4.1.3.1 Enquete-Kommission[1] C.4.1.3.2 Ausschuss für Bildung, Forschung & Technikfolgen-abschätzung1 C.4.1.4 Bund & Länder[1] C.4.1.4.1 Gemeinsame Wissenschaftskonferenz[1] C.4.1.4.2 Kultusministerkonferenz[1]

[4] Analysekategorie, die zur gemeinsamen Auswertung implizit und explizit solutionistischer Rechtfertigungen herangezogen wurde.
[5] Analysekategorie, die zur Quantifizierung der expliziten Kritik an der Cité der Solution ergänzt wurde.
[6] Im Zuge der Analyse wurden weitere Subkategorien für die jeweiligen im konfliktstehenden Cités angelegt (z. B. Staatsbürger vs. Solution), um fundiertere quantifizierende Aussagen über Existenzprüfungen treffen zu können.

		C.4.2 Wissenschaftliche Akteur:innen
		C.4.2.1 TAB[1]
		C.4.2.2 Wissenschaftsrat (WR)[1]
		C.4.2.3 Expertenkommission Forschung & Innovation (EFI)[1]
		C.4.2.4 Sonstige[1]
		C.4.3 Gewerkschaftliche Akteur:innen
		C.4.4 Verbände
		C.4.5 Wirtschaftliche Akteur:innen
		C.4.6 Sonstige
D. Sonstiges	D.1 Sonstiges	

Kategoriensystem & -definitionen

A. Analyserelevante Kategorien

Erläuterung: Block A umfasst Kategorien, die zur Strukturierung der Datengrundlage herangezogen werden, indem sie die zu analysierenden Textstellen der Dokumente als Segmente erfasst und für die spätere Ergebnisauswertung zitierfähige Aussagen festhält. Darüber hinaus werden ebenfalls analyserelevante Kategorien festgehalten, die Unklarheiten oder Ideen für neue Kategorien abbilden.

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung
A.1 Suchbegriffe	Automatisch codierte Sätze, die vordefinierte Suchbegriffe enthalten	[automatisiert codiert]		
A.2 Segmente	Segmente bilden die zu codierenden Sinneinheiten. Eine Sinneinheit bezeichnet den kleinsten inhaltlich abgeschlossenen Textabschnitt, der eine eigenständige Aussage oder Argumentation enthält und daher als Grundlage der Codierung dient. Sie umfasst typischerweise einen vollständigen Gedanken, der unabhängig von seiner formalen Länge in sich kohärent ist.	Satz als Minimaleinheit, häufiger Abschnitte, die die Rechtfertigung von KI thematisieren		
A.3 Zitierfähige Aussagen	Textabschnitte, die in besonderer Weise dazu geeignet sind, gewonnene Erkenntnisse zu belegen	Ausschnitte einzelner Segmente, die bspw. den Kern einer Cité widerspiegeln		
A.4 Unklarheiten	Segmente, bei denen die Zuordnung zu einer bestimmten Kategorie unklar erscheint	Unklarheit codieren und in Memo erläutern, anschließende Diskussion der Codierenden		
A.5 Neue Ideen	Vorschläge für induktiv zu bildende Kategorien	Neue Ideen codieren und in Memo erläutern, anschließende Diskussion der Codierenden		
A.6 Blumen am Wegesrand	Aspekte, die nicht von direkter Relevanz für die Beantwortung der Forschungsfragen sind, jedoch für den Gesamtzusammenhang der Arbeit von Interesse sind oder weiterführende Gedanken und/oder Interpretationsmöglichkeiten bieten	„Blumen“ codieren und in Memo erläutern, anschließende Diskussion der Codierenden		

B. Kontext

Erläuterung: Mit Hilfe der Kategorien in Block B wird analysiert, in welcher Form und in welchem Kontext KI im Dokument bzw. in der Analyseeinheit thematisiert wird. Es werden alle Segmente codiert, die sich mit irgendeiner Form der technischen Veränderung des Bildungswesens beschäftigen (von Lernen bis Administration).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
B.1 KI-Bezug	Ordnungskat.	Codierung von Subkategorien		
B.1.1 KI-Anwendung konkret	Ordnungskat.	Codierung von Subkategorien		… über „die KI“ gesprochen wird.
B.1.1.1 KI zum Lernen	KI-Anwendungen, die zum Lernerfolg beitragen sollen und den Lernenden dienen	KI-Anwendungen im Bereich des Lernens, die Studierenden oder Schüler:innen nutzen sollen	Individuelles Leistungsfeedback; individuelle Lernempfehlungen; „wird in Area9 Rhapsode der Lernpfad in Echtzeit nach der ausgewerteten Kompetenz und Selbstsicherheit des einzelnen lernenden sowie dessen ermittelten Wissenslücken laufend erstellt“ (LT-SN_894, 2021, S. 2 f.)	
B.1.1.1.1 Blended/E-Learning	Thematisierung von Blended oder E-Learning	Expliziter Bezug auf Blended- oder E-Learning Formate	„Blended learning bietet auch nach Auffassung des Deutschen Volkshochschulverbandes Möglichkeiten, die es stärker zu nutzen gilt.“ (BT_243a, 2019, S. 6)	
B.1.1.2 KI in der Lehre	KI-Anwendungen, die Dozierende unterstützen und die Lehre verbessern	KI-Anwendungen im Bereich der Lehre, die insbesondere Dozierende und Lehrer:innen unterstützen sollen	Automatisierte Lehrplanerstellung; adaptive Lehrunterstützung; „Auch die Lehrenden können durch eine Art ‚KI-Teaching-Assistant‘ begleitet werden, der ihnen bei der Qualitätsentwicklung ihrer Lehre hilft.“ (HFD_1201, 2021, S. 2)	

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
B.1.1.3 KI zur Bewertung	KI-Anwendungen, die zum Zwecke der Bewertung & Vorhersage eingesetzt werden	KI-Anwendungen im Bereich automatisierte Bewertung & Vorhersage, die sowohl Dozierende als auch Studierende unterstützen	Automatisierte Leistungsprognose; automatisierte Benotung; „Mit Hilfe dieser Daten können Lernfortschritte gemessen, zukünftige Leistungen prognostiziert (…) werden.“ (WR_1177, 2022, S. 38)	
B.1.1.4 KI zu Forschungszwecken	KI-Anwendungen, die in der Forschung eingesetzt werden	KI-Anwendungen im Bereich der Forschung, die im Forschungsprozess unterstützen sollen	Automatisierte Zusammenfassung eines Forschungsstandes; „Die Entwicklung von Künstlicher Intelligenz (…) wirken bereits disruptiv in der gesamten Forschungslandschaft.“ (GWK_1044, 2018, S. 395)	… Forschung an einer konkreten KI-Anwendung thematisiert wird, die nicht selbst Einfluss auf den Forschungsprozess nimmt.
B.1.1.5 KI in der Administration	KI-Anwendungen, die administrative Prozess erleichtern sollen	KI-Anwendungen im administrativen Bereich, die verwaltungstechnischen Aufwand reduzieren sollen	Reduktion von Routinetätigkeiten; Prognose von Studierendenzahlen; Automatisiertes Zulassungssystem; „(…) auf eine Beschleunigung der Bearbeitung von BAföG-Anträgen durch den Ausbau der Digitalisierung unter Nutzung künstlicher Intelligenz hinzuwirken.“ (BT_455c, 2022, S. 29)	… der Zweck des KI-Systems direkt den Lernerfolg beeinflussen soll (dann: Lernen).
B.1.2 KI-Anwendung diffus	Thematisierung diffuser KI-Anwendungen („die KI“)	Unspezifische Nennung von KI-Anwendung, i. S. v. „die KI“ oder „zahlreiche KI-Anwendungen“	Forderungen KI müsse Einzug in Klassenzimmer halten; „Verbesserung der Qualität, Leistungsfähigkeit und Wirksamkeit der Hochschulbildung durch den Einsatz von KI“ (GWK_1052, 2020, S.1)	… die Investition in KI (statt in eine/mehrere Anwendungen/Systeme) im Vordergrund steht.
B.1.3 Wissenserwerb über KI	Thematisierung des Wissenserwerbs über KI	Thematisierung von (mehr) Bildung und Wissen über KI; auch Forderungen des Erlernens des Umgangs	AI Literacy, KI in Curriculum aufnehmen; KI-Kompetenzen stärken; „sicherzustellen, dass Lehrpersonal bereits in der Ausbildung in ausreichendem Maße mit Methoden der Künstlichen Intelligenz vertraut und geschult wird.“ (BT_288a, 2020, S. 3).	… mehr Ressourcen für Forschung über KI gefordert werden (dann: Investitionen).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
B.1.4 Investition in KI	Thematisierung von Investitionen in KI	Investitionen (finanziell oder personell) in KI-Entwicklung oder Forschung	Ausrufung von 50 neuen KI-Professuren; Bereitstellung von 1 Mio. € für die Erforschung von KI-basierten Lernempfehlungen; „Außerdem soll die finanzielle Ausstattung der AlphaDekade verbessert werden.“ (BT_288b, 2021, S. 3)	
B.1.5 Forschung über KI	Thematisierung von Forschung über KI im Allgemeinen	Nennung Forschung über KI, die auf keinen Anwendungsfall spezialisiert ist	„Der Erfolg der KI-Forschung in den Anwendungsbereichen Arbeit und Bildung hängt ebenso wie in anderen Bereichen in besonderem Maße von einer fruchtbaren Interaktion zwischen Grundlagenforschung, problemorientierter Grundlagenforschung und unmittelbar anwendungsorientierter Forschung ab“ (BT_113b, 2020, S. 288)	… Forschung in konkreten Einsatzgebieten/Kontexten thematisiert wird, die nicht dem Bildungssektor zuzuordnen sind; … sowie, wenn über Forschungsförderung gesprochen wird (dann: Investitionen in KI); … und, wenn Allgemeinwissen oder Bildung über KI eines/einer jeden Einzelnen gesprochen wird (dann: Wissenserwerb über KI).
B.1.6 Sonstige	Sonstiger KI-Bezug	Thematisierung von KI, die keinem vorherigen Bezüge zugeordnet werden kann		
B.2 Verortung im Bildungssystem	Ordnungskategorie	Codierung von Subkategorien		
B.2.1 Elementarbereich	KI in Kitas & Vorschulen	Thematisierung von KI im Vorschul- und Kindergartenbereich	„Für den Kita-Bereich empfiehlt die SWK, ‚Digitale Medienbildung‘ und elementarinformatische Bildung in allen Bildungs- und Orientierungsplänen der Bundesländer zu verankern“ (GEW_1314_2022, S. 1)	

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
B.2.2 Primar- & Sekundarbereich	KI in der Grundschule (Klassen 1-4) oder der Sekundarstufen (I & II, Klassen 5-13)	Thematisierung von KI an Grundschulen, Haupt-, Realschulen oder Gymnasien	„Eine Auseinandersetzung mit den Chancen und Risiken von KI muss bereits in den Grundschulen erfolgen.“; „Schüler:innen des Gymnasiums beschäftigen sich mit der ‚Aktionswoche KI‘.“; „In der Grundschule sollen informatische Inhalte im Rahmen des Sachkundeunterrichts behandelt werden, ab der Sekundarstufe I in einem Pflichtfach Informatik.“ (GEW_1314_2022, S. 1)	
B.2.3 Tertiärbereich	KI im Hochschulsystem	Thematisierung von KI im Bereich von Hochschulen, Universitäten, Fachhochschulen, Berufsakademien und/oder Fachakademien	Automatisierte Studienplatzvergabe; KI-Studiengänge; KI-Professuren; „Um in der Vielfalt und Breite des Hochschulsystems wirksame Effekte in Studium und Lehre zu erreichen (…)“ (GWK_1052, 2020, S. 1)	
B.2.4 Quartärbereich	KI im Weiterbildungsbereich	Thematisierung von KI im Bereich privater oder beruflicher Weiter- und Fortbildung; Erwachsenenbildung	„Neuste KI-Technologien erfordern Umschulungen und Weiterbildungen im Arbeitsmarkt.“; „In den beiden anderen Anträgen geht es um Lernen und Weiterbildung in der KI, um junge Menschen und auch Menschen, die schon im Berufsleben sind, entsprechend weiterzuqualifizieren.“ (BT_244c, 2019, S. 182)	
B.2.5 Unspezifizierter Bildungsbereich	KI im Bildungskontext, der keinem spezifischen Bereich zuzuordnen ist	Thematisierung von KI im Bereich des Bildungssektors, ohne dass eine Zuordnung zu Schulen, Universitäten oder anderen Bereichen vorgenommen wird.	„Forschung und Entwicklung neuer innovativer Lernmethoden zu gewährleisten, in dem sie Unternehmen unterstützt, die sich gezielt mit der Entwicklung und Anwendung von KI in Verbindung mit Bildung widmen.“ (BT_288a, 2020, S. 4)	

C. Rechtfertigungen

Erläuterung: Mit Hilfe der Kategorien in Block C wird analysiert, welche Rechtfertigungen für oder gegen den Einsatz von KI-Anwendungen im deutschen Bildungssystem geäußert werden. Darunter fallen Begründungen, Erklärungen, Argumente, Gründe oder Rechtfertigungen für oder gegen den Einsatz von KI sowie Forderungen oder Kritik nach/an KI-Anwendungen.

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.1 Cité	Ordnungskat.	Codierung von Subkategorien		
C.1.1 Inspiration	Rechtfertigung aus Sicht der Cité der Inspiration	Rechtfertigungen, die Genialität, Kreativität, Übermenschliches oder Visionäres der KI in den Vordergrund rücken (s. weiterführende Hinweise „zum Verständnis der Cités)	Sakralisierung einzelner KI-Entwickler:innen (z. B. Steve Jobs); Betonung des Übermenschlichen; „Es geht darum, dass man kreative Ideen umsetzen kann, die sie so im normalen Unterricht noch nie umsetzen konnten." (LT-TH_850, 2021, S. 3551)	… die Problemlösungsfähigkeit von KI im Vordergrund steht (dann: Solution).
C.1.2 Haus	Rechtfertigung aus Sicht der häuslichen Cité	Rechtfertigungen, die auf den Erhalt oder Abbau bestehender (sozialer) Ordnungen und Hierarchien rekurrieren, Tradition, Bewährtes und die Stabilität vertrauter gemeinschaftlicher Ordnungen hervorheben (s. weiterführende Hinweise „zum Verständnis der Cités)	Gefährdungspotenzial für bestehende Ordnungen; Bewahrung des Bestehenden; Präferenz von Uniformität, sozialer Kohäsion; Vermeidung der Offenlegung individueller (Leistungs-)Unterschiede; Betonung partikularistischer Gemeinschaft; „Die SWK zeichnet das Szenario eines auf Algorithmen gestützten Unterrichts, der tiefgreifende Veränderungen und Eingriffe (…) in das Professionsverständnis der Lehrenden und den Umgang zwischen Lernenden und Lehrenden hätte." (GEW_1314, 2022, S. 2)	
C.1.3 Meinung	Rechtfertigung aus Sicht der Cité der Meinung	Rechtfertigungen, die sich auf öffentliche Zustimmung, Mehrheitsmeinungen, Ruhm, Reputation, mediale Aufmerksamkeit oder prominente Vorbilder beziehen (s. weiterführende Hinweise „zum Verständnis der Cités)	Referenzen auf Bevölkerungsumfragen; Heranziehung anderer Länder als Paradebeispiele; Bezug zu Wähler:innenpotenzialen; „dafür zu werben, dass in der Bevölkerung insgesamt eine breite Akzeptanz künstlicher Intelligenz als Lernhilfsmittel entsteht." (BT_288a, 2020, S. 4)	… andere Länder i. S. eines wirtschaftlichen Wettbewerbs genannt werden (dann: Markt).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.1.4 Staatsbürger	Rechtfertigung aus Sicht der staatsbürgerlichen Cité	Rechtfertigungen, die Gemeinwohl, Solidarität, Gleichheit, Nichtdiskriminierung, demokratische Verfahren oder Rechtsstaatlichkeit betonen (s. weiterführende Hinweise „zum Verständnis der Cités)	Ethische Prinzipien der KI-Entwicklung; Menschenzentriertheit; Diskriminierung; Chancengleichheit; Erfüllung des Humboldtschen Bildungsideals (Entwicklung mündiger Bürger:innen); Solidarität, antipartikularistisch; „Gesellschaftliche Ungleichheit und Diskriminierung von Individuen (…) darf nicht fortgeschrieben oder gar verstärkt werden, indem in die Programmierung oder durch die Trainingsdaten tendenziöse Vorfestlegungen oder Diskriminierungen einfließen.“ (BT_162b, 2019, S. 16)	
C.1.5 Industrie	Rechtfertigung aus Sicht der industriellen Cité	Rechtfertigungen, die Innovation und technischen Fortschritt, Effizienz- & Effektivitätsgewinne, Standardisierung, technische Kontrolle oder Planbarkeit in den Vordergrund rücken (s. weiterführende Hinweise „zum Verständnis der Cités)	Effizienz- & Effektivitätsgewinn durch KI-Einsatz; Ressourceneinsparung; Versuch der Metrisierung; Formulierung von Zielvorgaben; Technologischer Fortschritt (zum effizienteren Wissenserwerb); Planungssicherheit; Ungleichheit wird in Kauf genommen; Qualifizierungsmaßnahmen; „Rhapsode bestimmt in Echtzeit, worauf mehr Gewichtung gelegt, was korrigiert werden muss und wo Zeit gespart werden kann, indem redundante Inhalte vermieden werden.“ (LT-SN_894, 2021, S. 2)	… der ökonomische Gewinn im Vordergrund steht (dann: Markt).
C.1.6 Markt	Rechtfertigung aus Sicht der Cité des Marktes	Rechtfertigungen, die Wettbewerb, Konkurrenzfähigkeit des Staates und des Individuums (i. S. des Individualismus als persönlichen Marktvorteil) thematisieren (s. weiterführende Hinweise „zum Verständnis der Cités)	Internationale Wettbewerbsfähigkeit (soll Bildungserfolg erhöhen); Vergleich KI-Entwicklungsstand in anderen Ländern; ökonomisches Potenzial; Betonung des Individualismus (z. B. individuelle Lernpfade); Qualifizierung von Fachkräften (als Wettbewerbsvorteil); „Eine Stärkung der Erforschung und Anwendung von KI in Deutschland leistet einen wichtigen Beitrag zur Stärkung der Wettbewerbsfähigkeit dieses Bereiches in Deutschland und zugleich in Europa.“ (BT_162b, 2019, S. 18)	… individuelle Angebote mit dem expliziten Ziel der Chancengleichheit („KI soll allen zugutekommen“) im Vordergrund stehen (dann: Staatsbürger).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.1.7 Projekt	Rechtfertigung aus Sicht der projektbasierten Cité	Rechtfertigungen, die Flexibilität, Netzwerkbildung und flache Hierarchien, situative Problemlösung oder projektorientiertes Handeln in den Mittelpunkt stellen (s. weiterführende Hinweise „zum Verständnis der Cités)	KI-Anwendungen als Pilotprojekte; Kooperationen zur Erreichung von Zielen oder zur Lösung von Problemen; „Da bei der Bildung mit digitalen Medien immer das Zusammenspiel von technischer Ausstattung, der Verfügbarkeit entsprechender Angebote und der Kompetenz der Lehrenden wichtig ist, sollten Pilotprojekte exemplarisch alle drei Aspekte adressieren und ihr Zusammenspiel optimieren.“ (TAB_1335, 2020, S. 22)	… technologischer Fortschritt thematisiert wird (dann: Industrie).
C.1.8 Ökologie	Rechtfertigung aus Sicht der grünen Cité	Rechtfertigungen, die Nachhaltigkeit, Umwelt- und Ressourcenschutz, Gesundheit, Zukunftsfolgen oder ökologische Risiken betonen (s. weiterführende Hinweise „zum Verständnis der Cités)	Energiebedarf von KI-Systemen oder der Verarbeitung großer Datenmengen; Folgen für zukünftige Generationen; „Digitalisierung und Ökologie gehören zusammen, und deswegen sage ich Ihnen auch den sozialdemokratischen Ansatz: Wir sprechen von nachhaltiger Digitalisierung, weil „digital“ nicht automatisch „öko“ ist.“ (BT_251f, 2020, S. 58)	… das Allgemeinwohl der Gesellschaft thematisiert wird (dann: Staatsbürger).
C.1.9 Solution	Rechtfertigung aus Sicht der Cité der Solution	Rechtfertigungen, die soziale Probleme als technisch lösbar darstellen und KI als beste oder einzige Option für schnelle Lösungen präsentieren (s. weiterführende Hinweise „zum Verständnis der Cités)	KI-Anwendungen als Problemlöser; Beispiel, wenn Videoidentifikation durch KI ersetzt würde: „Eine einfache Videoidentifikation per Smartphone, wie sie für die Einrichtung eines Bankkontos durchaus üblich ist, steht nicht zur Verfügung. Das wäre ja zu einfach." (BT_251e, 2019, S. 9); „Damit sind Anwendungen gemeint, die explizit gesellschaftliche Probleme lösen möchten“ (BT_113b, 2020, S. 201)	… KI als Übermenschliches dargestellt wird (dann: Inspiration).
C.1.10 Andere Form der Rechtfertigung	Offene Kategorie der Rechtfertigung	Rechtfertigungen, die keiner der anderen neun Rechtfertigungsordnungen zugeschrieben werden können.		

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.1.11 Keine Rechtfertigung	Keine Rechtfertigung einer Forderung der Auseinandersetzung mit/Investition in KI etc.	Äußerungen, die den Einsatz von KI, die Investition in oder den Wissenserwerb über KI fordern, aber dafür keinerlei Rechtfertigung (Definition s. Glossar) liefern.	„Verbesserung der Hochschulbildung durch den Einsatz von KI Hochschulen oder Hochschulverbünde können einmalig beispielsweise bei der Gestaltung (…) von KI-basierten kollaborativen Lehr- und Lernnetzwerken von Studierenden (…), bei KI-gestützter Lernprozessunterstützung, KI-gestützter Kurs- und Modul(weiter)entwicklung sowie bei der Entwicklung KI-gestützter Studienberatung für in- und ausländische Studieninteressierte und Maßnahmen KI-gestützter (Selbst-)Organisation des Studiums gefördert werden.“ (GWK_10.12.2020_1052, S. 2)	
C.1.12 Solution (implizit & explizit)	Analysekategorie	Automatisiert erstellte Kategorie auf Basis der Kombination aller implizit (C.1.11) und explizit solutionistischen Rechtfertigungen (C.1.9)		
C.1.13 Anti-Solution	Analysekategorie: kritische Äußerungen, die sich gegen solutionistische Sichtweisen richten	Rechtfertigungen, die solutionistische Sichtweisen explizit kritisieren; i. d. R. geschieht dies aus Sicht einer anderen Cité	Abstreiten, dass KI-Anwendungen als gesellschaftliche Problemlöser jedweder Herausforderungen dienen können; „Das neue Buzzword ist KI, und deswegen muss KI diese Woche in jedem Antrag der FDP vorkommen, Sinnhaftigkeit second.“ (BT_244c, 2019, S. 185)	
C.2 Prüfungen	Ordnungskat.	Codierung von Subkategorien		
C.2.1 Wahrheitsprüfung	Selbstbestätigende Aussagen	Tautologische, selbstreferenzielle oder normstabilisierende Rechtfertigungen, die frei von Kritik sind	„Die Bundesregierung misst der Förderung des wissenschaftlichen Nachwuchses im Bereich KI besondere Bedeutung zu und wird noch in diesem Jahr ein entsprechendes Programm für die genannte Zielgruppe auf den Weg bringen.“ (BT_162b, 2019, S. 11)	… Herausforderungen/Regulierungsbedarfe des KI-Einsatzes thematisiert werden (dann: Realitätsprüfung).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.2.2 Realitätsprüfung	Bewährungsproben einer Cité	Herausforderungen, die der KI-Einsatz für eine Cité mit sich bringt; häufig in Verbindung mit reformistischer Kritik	„Sie weist darauf hin, dass auch bei der Diskussion um Bildung in einer digitalen Gesellschaft ganzheitliche Bildung das Ziel sein müsse. Man sei daher bei dem Thema KI in der Schule mindestens skeptisch. Dies wäre ein Thema, über das man beispielsweise einen umfassenden TAB-Bericht anfertigen lassen könnte. Da zum jetzigen Zeitpunkt noch zu viele Fragen offen seien, Achtsamkeit gefragt sei und man den Prozess zunächst kritisch beobachten müsse, lehne man den Antrag der Fraktion der FDP ab.“ (BT_244b, 2020, S. 8)	… Kritik aus der Sicht einer Cité an einer anderen Cité geäußert wird (dann: Existenzprüfung).
C.2.3 Existenzprüfung/Konflikt	Konflikte zw. min. zwei Cités	Rechtfertigungen, die im Konflikt zueinanderstehen bzw. Kritik, die an einer anderen Cité geäußert wird; mitunter verbunden mit radikaler Kritik des KI-Einsatzes	Rechtfertigungen, die Effizienzgewinne (Industrie) zulasten der Chancengleichheit (Staatsbürger) kritisieren; „Die Fraktion DIE LINKE. betont, dass es bei der Bildung nicht darum gehe, wie man mit wenigen Ressourcen viel Stoff möglichst schnell in die Köpfe bekomme. Vielmehr gehe es darum, Bildung als sozialen Prozess zu gestalten.“ (BT_244b, 2020, S. 7)	… zwei Cités gemeinsam thematisiert werden, ohne dass ein Konflikt zw. ihnen ersichtlich wird (dann: Kompromisse).
C.2.4 Kompromisse	Gemeinsame Rechtfertigung min. zweier Cités	Rechtfertigungen, die die moralische Grundlage zweier (oder mehr) Cités verbinden, um den jeweiligen Standpunkt zu stärken. Cités stehen dabei nicht im Konflikt, sondern ergänzen sich, um Rechtfertigung zu sichern.	„Für eine noch effizientere und bürgerfreundlichere Verwaltung soll die Landesverwaltung bundesweit Pionier beim Einsatz von KI und der besseren Nutzung von Daten mit den auf europäischer Ebene etablierten Datenschutzstandards und -vorgaben werden“ (LT-SH_670, 2019 S. 58)	… zwei Cités gemeinsam thematisiert werden, die in einem Konflikt zueinander stehen bzw. eine Cité Kritik an einer anderen Cité äußert (dann: Konflikte).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.3 Richtung der Rechtfertigung	Ordnungskategorie	Codierung von Subkategorien		
C.3.1 Pro-KI	Rechtfertigungen für den Einsatz von KI, i. S. positiver (erwarteter) Folgen	Argumente, Rechtfertigungen, Forderungen oder Gründe für den Einsatz von KI-Systemen oder die Investition in KI; positive Folgen/positiver Zweck des KI-Einsatzes	„Das Bildungsland Deutschland und das Innovationsland Deutschland gehen Hand in Hand (…) in eine wettbewerbsfähige Zukunft, (…) eine Zukunft, in der Kreativität und Empathie der Kern eines neuen Bildungsversprechens sind.“ (BT_244d, 2021, S. 50); „Durch den (…) Einsatz von KI im Bildungsbereich (…) eröffnen sich Möglichkeiten, Lernangebote in Zukunft individueller zu gestalten und auf konkrete Bedürfnisse der Lernenden und Lehrenden einzugehen.“ (BT_113b, 2020, S. 297)	… neben der Befürwortung des KI-Einsatzes auch Hindernisse thematisiert werden (dann: Abwägend Positiv).
C.3.2 Abwägend Positiv	Resümierende Rechtfertigung für den Einsatz von KI nach Abwägung von Vor- & Nachteilen	Argumente, Rechtfertigungen, Forderungen oder Gründe für oder gegen den Einsatz von KI-Systemen, die Investition in oder den Wissenserwerb über KI oder Folgen bzw. Zweck des KI-Einsatzes, wobei Argumente für einen KI-Einsatz überwiegen	„Aufgabe und Ziel der Politik muss es sein, die Entwicklung und den Einsatz Künstlicher Intelligenz (KI) so zu gestalten, dass die vielfältigen Chancen der Technologie für das Gemeinwohl nutzbar gemacht werden und die Rechte der Bürgerinnen und Bürger gewahrt werden.“ (BT_162b, 2019, S. 1)	… keine eindeutige Tendenz erkennbar ist, ob der KI-Einsatz befürwortet oder abgelehnt wird (dann: keine Valenz).
C.3.3 Keine Valenz	Neutral beschreibende Rechtfertigung des Einsatzes oder der Beschäftigung mit KI	Sachliche Beschreibung des Einsatzes von KI-Systemen, der Investition in oder des Wissenserwerbs über KI ohne, dass positive oder negative Folgen bzw. positive/negative Zwecke des KI-Einsatzes eindeutig hervorgehoben werden	„Angesichts der bereits initiierten Forschungsvorhaben, die sowohl die praktische Entwicklung von Learning Analytics und deren Einsatz in verschiedenen Bildungskontexten und insbesondere im Hochschulbereich als auch die Auseinandersetzung mit den Voraussetzungen und Auswirkungen umfassen, erscheint eine Vertiefung des Themas im Zuge der parlamentarischen Technikfolgenabschätzung zum jetzigen Zeitpunkt noch zu früh.“ (TAB_1333, 2021, S. 7)	… (implizit) die Befürwortung oder Ablehnung präferiert wird (dann: Abwägend Positiv/Negativ).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.3.4 Abwägend Negativ	Resümierende Rechtfertigung gegen den Einsatz von KI nach Abwägung von Vor- & Nachteilen	Argumente, Rechtfertigungen, Forderungen oder Gründe für oder gegen den Einsatz von KI-Systemen, die Investition in oder den Wissenserwerb über KI oder positive/negative Folgen bzw. positiver/negativer Zweck des KI-Einsatzes, wobei die Argumente gegen einen KI-Einsatz überwiegen	„Es gibt jedoch auch große und begründete Bedenken gegen eine überstürzte Einführung an Schulen — sowohl aus pädagogischer Sicht als auch im Hinblick auf Datenschutz. Aus pädagogischer Sicht müssen wir aufpassen, dass das Sammeln von Daten kein Selbstzweck wird. Learning Analytics ist pädagogisch erst dann relevant, wenn aus der Unmenge Daten auch wertvolle Aussagen entstehen; denn Lernverhalten ist ein Bildungsprozess, der nicht vollständig statistisch erfasst werden kann.“ (BT_244c, 2019, S. 232)	… keine eindeutige Tendenz erkennbar ist, ob der KI-Einsatz befürwortet oder abgelehnt wird (dann: Keine Valenz).
C.3.2 Contra-KI	Rechtfertigungen gegen den Einsatz von KI oder Nennung von negativen Folgen	Argumente, Rechtfertigungen oder Gründe gegen den bzw. Kritik am Einsatz von KI-Systemen oder der Investition in KI oder des Wissenserwerbs über KI; Risiken, (erwartete) negative Folgen oder Schäden	„Learning Analytics reduziert unsere Schüler und auch Lehrkräfte auf reine Datenlieferanten. Alle erzeugen Daten. Lehrer sollen ihren Unterricht mit Cloud-Tools planen und in einer zentralen Cloud speichern, auf die weiß Gott wer nachher Zugriff hat. Wie weit sind wir dann eigentlich noch davon entfernt, lediglich Wirte einer allmächtigen Datenmatrix zu sein? (…) Das ist nicht unsere Welt, und das ist auch nicht mein Bildungsideal. Wir in der AfD vertreten das Bildungsideal von Humboldt: Schüler sollen mündige Erwachsene werden, die ganzheitlich gebildet sind. Deswegen werden wir Ihren Antrag ablehnen.“ (BT_244c, 2019, S. 183)	… neben der Kritik des KI-Einsatzes auch Potenziale thematisiert werden (dann: Abwägend Negativ).

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.4 Rechtfertigende:r Akteur:in	Ordnungskategorie	Codierung von Subkategorien		
C.4.1 Politische Akteur:innen	Ordnungskategorie	Codierung von Subkategorien		
C.4.1.1 Parteien	Äußerungen im Namen einer Partei oder Vertreter:innen selbiger	Codierung folgender Parteien: - C.4.1.1.1 Die Linke - C.4.1.1.2 Bündnis 90/Die Grünen - C.4.1.1.3 SPD - C.4.1.1.4 CDU/CSU - C.4.1.1.5 FDP - C.4.1.1.6 AfD - C.4.1.1.7 Andere Partei	„Ein Verständnis von Bildungspolitik, die gleichzeitig Wirtschaftsförderung betreibe, sei abzulehnen. Nach Auffassung der Fraktion DIE LINKE. gehe es in der Schule, (…) ursprünglich um Bildung, die im Zusammenhang mit Freiheit und Unabhängigkeit stehe.“ (BT_195b, 2019, S. 10)	… von Äußerungen „der Bundesregierung“ als Kollektiv gesprochen wird.
C.4.1.2 Regierungen	Regierungsäußerungen	Äußerungen, die von „der“ Bundes- oder Landesregierung als Kollektiv erfolgen	„Die Bundesregierung misst der Förderung des wissenschaftlichen Nachwuchses im Bereich KI besondere Bedeutung zu“ (BT_162b, 2019, S. 11); „Gleichwohl unterstützt die Landesregierung durch geeignete Formate den Austausch von Akteurinnen und Akteuren im Bereich der Künstlichen Intelligenz“ (LT-NW_895, 2021, S. 171)	… Vertreter:innen von Parteien auftreten, die zwar an der Regierung beteiligt sind, jedoch nicht explizit im Namen der Regierung sprechen (dann: Parteien).
C.4.1.3 Parlamentarische Gremien	Äußerungen parlamentarischer Gremien	Codierung folgender Gremien: - C.4.1.3.1 Enquete-Kommission - C.4.1.3.1 Ausschuss für Bildung, Forschung und Technikfolgenabschätzung	„Die Einstufung von KI-Systemen in Kritikalitätsstufen war in der Enquete-Kommission allgemein und in der Projektgruppe umstritten.“ (BT_113b, 2020, S. 287); „Der Ausschuss fordert die Bundesregierung auf“ (BT_455c, 2022, S. 29)	

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.4.1.4 Bund & Länder	Zusammenschlüsse /Gremien aus Bund & Ländern	Codierung folgender Zusammenschlüsse: - C.4.1.4.1 Gemeinsame Wissenschaftskonferenz - C.4.1.4.2 Kultusministerkonferenz	„Der Vorsitzende der Gemeinsamen Wissenschaftskonferenz (…) erklärt“ (GWK_1058, 2021, S. 1); „Für Kinder und Jugendliche ist es unabdingbar frühzeitig zu lernen, adäquat und reflektiert mit digitalen Medien umzugehen, wie es der Kompetenzrahmen der KMK-Strategie aufzeigt“ (KMK_1076, 2021, S. 9)	
C.4.2 Wissenschaftliche Akteur:innen	Wissenschaftler:innen oder (Vertreter:innen) wissenschaftliche(r) Institutionen, wie Universitäten oder unabhängige Forschungsinstitute	Codierung folgender wiss. Einrichtungen - C.4.2.1 Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB) - C.4.2.2 Wissenschaftsrat (WR) - C.4.2.3 Expertenkommission Forschung & Innovation (EFI) - C.4.2.4 Sonstige	„Das TAB hat sich schon in den 2000er Jahren im Rahmen eines umfangreichen Monitorings zum »eLearning« mit den Aspekten des Einsatzes von digitalen Medien in den verschiedenen Bildungsbereichen in grundsätzlicher Hinsicht befasst.“ (TAB_1335, 2020, S. 31); „erklärte die Erziehungswissenschaftlerin, die an der Helmut-Schmidt-Universität der Bundeswehr in Hamburg forscht.“ (GEW_1239, 2019, S. 1)	
C.4.3 Gewerkschaftliche Akteur:innen	Gewerkschaftler:innen oder (Vertreter:innen) gewerkschaftliche(r) Institutionen	Akteur:innen, die im Namen einer Gewerkschaft auftreten, hier i. d. R. der GEW	„Die GEW stellt sich den Risiken und Chancen einer durch Automatisierung, Algorithmen und künstlicher Intelligenz beeinflussten Gesellschaft.“ (GEW_1338, 2021, S. 5)	
C.4.4 Verbände	Verbände oder Vertreter:innen von Verbänden und zivilgesellschaftlichen Organisationen	Akteur:innen, die im Namen eines Verbandes auftreten, die keine Gewerkschaften sind, hier i. d. R. HFD; Ein Verband „bezeichnet den Zusammenschluss von Personen mit gemeinsamen Interessen zur Verfolgung gemeinsamer Ziele“ (Schubert & Klein, 2020)	„Dazu die HFD-Arbeitspapiere (…) und das entsprechende Themendossier des HFD.“ (HFD_1211, 2021, S. 3)	

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung Wird nicht codiert, wenn …
C.4.5 Wirtschaftliche Akteur:innen	Wirtschaftliche Akteur:innen oder (Vertreter:innen) wirtschaftliche(r) Institutionen, wie kommerzieller Unternehmen	Direkt oder indirekt wiedergegebene Aussagen von Vertreter:innen privatwirtschaftlicher Unternehmen	„Auf Nachfrage schränkt die Geschäftsführung von Retorio derartig weitreichende Leistungsversprechen ein und verweist darauf, dass Auswahlentscheidungen nie 100%ig neutral sein können." (TAB_1327, 2021, S. 5)	… wenn über Unternehmen gesprochen wird, ohne deren Standpunkte darzulegen.
C.4.6 Sonstige	Akteur:innen, die keiner der vorherigen Kategorien zugeordnet werden können	Akteur:innen, deren Funktion unklar bleibt		

D. Sonstiges

Sonstige Aspekte, die keinen bestehenden Kategorien zugeordnet werden können

Kategorie	Definition	Anwendung	Beispiel(e)	Abgrenzung
D.1 Sonstiges	Neuauftauchende Aspekte, die keiner bestehenden Kategorie zugeordnet werden können, aber von Relevanz zur Beantwortung (einer) der Forschungsfragen sind.	In der Regel Einzelnennungen (oder Aspekte, die weniger als vier Mal vorkommen); Kommen Aspekte mehrfach vor, wird gemeinschaftlich über die Ergänzung neuer Kategorien gesprochen; Besonders relevante Aspekte, für die eigener Code sinnvoll erscheint, werden direkt als Subcode eines „Sonstiges"-Codes eingefügt, wobei der eigene Name in Klammern ergänzt wird; neu gebildete Subcodes erhalten (vorläufige) Definition im Code-Memo sowie Farbmarkierung „braun"; Neue Subcodes werden über Basecamp kommuniziert und im Projekt-Logbuch festgehalten.		

Zum Verständnis der Cités

Cités – oder in der deutschen Übersetzung auch als *Rechtfertigungsordnungen*, *Polis* oder *Gemeinwesen* bezeichnet – werden verstanden als unterschiedliche **Gerechtigkeitsvorstellungen** innerhalb einer Gesellschaft, die von zentraler Bedeutung in sozialen Konflikten sind, da sie als moralische Grundsätze und somit als **rechtfertigende Gründe für das eigene Handeln herangezogen** werden (Bogusz, 2010; Forst, 2015). Cités sollen daher nicht selbst als Rechtfertigungen verstanden werden, sondern vielmehr als Äquivalenzprinzipien, die eine (moralische) Grundlage einer Rechtfertigung bilden.

Cité der Inspiration

Die Cité der Inspiration, verschiedentlich auch als Welt der Erleuchtung bezeichnet (Hartmann, 2008), orientiert sich am gottgegebenen und scheint insofern religiös motiviert (Forst, 2015). Demnach werden Hochmut und Habgier ebenso wie das Streben nach Ruhm[1] gänzlich abgelehnt und anstelle dessen **Demut** als Tugend geschätzt. Zur Erreichung der Inspiration, die von Boltanski und Thévenot (2018 [1991]) gewissermaßen als esoterischer und berauschender Zustand beschrieben wird, bedarf es den Autoren zufolge asketischer Enthaltsamkeit, insofern als dass öffentliche Aufmerksamkeit der **Gedankenfreiheit** im Wege stehe und die Anerkennung und Meinung anderer irrelevant oder gar hinderlich erscheine. In der moderneren Fassung kann die Cité der Inspiration vornehmlich als Welt der **Künstler:innen** begriffen werden, welche Raum für **Kreativität** bietet und großen Wert auf **Ästhetik** legt (Gadinger & Yildiz, 2012). Das Streben nach Originalität mündet in einem **Enthusiasmus**, der eine rationale Haltung vermissen lässt – und gar erfordert, diese wissentlich aufzugeben (Boltanski & Thévenot, 1991/2018, S. 222 ff.). Aufgrund dieser kreativen Ausrichtung erscheint die Cité der Inspiration gewissermaßen phantastisch und sehr impulsiv und ist somit die fluktuativste und unbeständigste Cité (Bogusz, 2010, S. 53).

Übertragen auf die konkrete Betrachtung der Entwicklung und Rechtfertigung modernster Technologien spiegelt sich die Cité der Inspiration in der Bewunderung von **Genialität** und der „zuweilen guruartige[n] Verehrung ‚gnadenbegabter' Entrepreneure (man denke nur an Steve Jobs)" (Nachtwey & Seidl, 2021, S. 248) wider. Vorstellbar ist jedoch auch, dass aus dieser Perspektive heraus die KI selbst gewissermaßen sakralisiert wird, als dass sie als Heilsbringer zur Erreichung höherer Zwecke symbolisiert wird.

Die häusliche Cité

Durch Wertschätzung von **Gewohnheit und Traditionen, Vertrauen, Verantwortungsgefühl und Respekt** sowie die Ablehnung von Neid, Schwätzertum und Verrat zeichnet sich die häusliche Cité aus (Bogusz, 2010; Boltanski & Thévenot, 1991/2018). Gleichzeitig werden die Akteurskonstellationen in dieser Welt stark durch eine als natürlich gegeben angenommene **Hierarchie** geprägt, wie sie in familiären Kontexten zu finden ist und persönliche Beziehungen und autoritäre Abhängigkeitsverhältnisse erzeugt (Forst, 2015; Gadinger & Yildiz, 2012; Hartmann, 2008; Hildebrandt, 2018). Dabei muss die Argumentation aus Sicht der häuslichen Cité jedoch nicht immer familiär geprägt sein, sondern kann sich auch in den Hierarchiestrukturen tradierter Unternehmen wiederfinden. Im Gegensatz zur Cité der Inspiration gilt daher nicht das „Geniale" oder das „Phantastische" als erstrebenswert, sondern das Tradierte. Häufig wird dabei Bezug auf das Handeln solcher Akteur:innen genommen, die sich bereits verdient gemacht haben und insofern in der Hierarchiekette oben stehen und als Autoritäten Respekt genießen (Boltanski & Thévenot, 1991/2018). Im Bildungssektor kommt die häusliche Cité in Form einer sozialen Kohäsion zum Ausdruck, die das Glück einer **partikularistischen Gemeinschaft** in

[1] Insofern steht die Cité der Inspiration in einem ständigen Spannungsverhältnis mit der Cité der öffentlichen Meinung, die sich an öffentlicher Anerkennung orientiert (Boltanski & Thévenot, 1991/2018).

den Fokus stellt. **Uniformität** gilt als erstrebenswert, denn individuelle (Leistungs-)Unterschiede sollen möglichst nicht aufgedeckt werden (Imdorf, 2011).

Vor dem Hintergrund des Traditionsbewusstseins ist davon auszugehen, dass eine KI-Einführung kritisch gesehen wird und insbesondere dann abgelehnt wird, wenn sie droht, bestehende Ordnungen zu gefährden. So kann eine Bestätigung des Status quo als Begründung aus Sicht der häuslichen Cité herangezogen werden.

Die Cité der (öffentlichen) Meinung

Ähnlich wie die Cité der Inspiration ist auch die Cité der öffentlichen Meinung stark fluktuativ. Wiederum gelten hier in Abgrenzung zur inspirierten Welt jedoch **Ruhm und Anerkennung** als erstrebenswerte Ziele, weshalb Hartmann (2008) auch von einer an **Reputation** orientierten Ordnung oder Gadinger und Yildiz (2012) von einer Cité des Renommées sprechen. Diese wird maßgeblich von den Medien mitgeprägt, sodass Rechtfertigungen auf (der Annahme über) **Mehrheitsmeinungen** konstruiert werden und den Aussagen sogenannter Meinungsführer besondere Beachtung geschenkt wird. Eigene Interessen werden insofern hintenangestellt, als dass das Streben nach öffentlicher Anerkennung und Wertschätzung in einem Abhängigkeitsverhältnis zu der Meinung anderer mündet (Boltanski & Chiapello, 2001; Boltanski & Thévenot, 1991/2018).

Plausible Rechtfertigungen für oder gegen den Einsatz von KI aus Sicht dieser Cité könnten daher beispielsweise von Politikern geäußert werden, die sich an einem größtmöglichen Wähler:innenpotenzial orientieren. Auch der Bezug zu anderen berühmten Persönlichkeiten oder „erfolgreichen" Staaten kann als Indikator aufgefasst werden (Boltanski & Thévenot, 1991/2018, 247 f.).

Die staatsbürgerliche Cité

In der staatsbürgerlichen, bürgerweltlichen oder auch zivilgesellschaftlichen Cité stehen wenig überraschend das **Kollektiv** und das klassische Gemeinwohl im Vordergrund. Demnach werden auch hier individuelle Interessen zurückgestellt, anders als bei der Cité der Meinung jedoch nicht zum Zwecke der Erlangung äußerer Anerkennung, sondern zugunsten von **Solidarität** und kollektiven Interessen (Forst, 2015; Gadinger & Yildiz, 2012). In der Konsequenz sollen **Individualismus und Partikularismus vermieden** werden und **Teilhabe und Mitbestimmung** spielen wichtige Rollen in der demokratischen Entscheidungsfindung – Willkür wird hingegen verurteilt (Bogusz, 2010). In diesem Sinne wird zum einen ein Humboldtsches Bildungsideal verfolgt, das mündige Bürger:innen heranwachsen lässt, und zum anderen der **Chancengleichheit** ein hoher Wert zugerechnet. Benachteiligungen jeglicher Art werden abgelehnt (Imdorf, 2011).

Demzufolge tritt die häufig mit algorithmischen Entscheidungen in Verbindung gebrachte Kritik der (Verschärfung von) Ungleichheit und Diskriminierung in dieser Cité besonders deutlich zutage (Peter, 2015, S. 183). Eine besondere Rolle kommt außerdem der **Legalität**, verstanden als Rechtmäßigkeit, zu. Der Gemeinwille soll möglichst in Form von Gesetzen und Beschlüssen festgeschrieben werden, sodass individuelle Abweichungen und die Verfolgung egoistischer Interessen erschwert werden (Knoll, 2012, S. 69). Außerdem wird eine demokratische Repräsentativität des Kollektivs gefordert, die ebenfalls festen Gesetzesgrundlagen unterliegt. Ein:e Repräsentant:in des Kollektivs besitzt zum einen die höchste Größe, zum anderen ist diese:r jedoch gegenüber dem Kollektiv auch rechenschaftspflichtig (Boltanski & Thévenot, 1991/2018, 261 f.). Ein Bezug zu ethischen Richtlinien der KI-Entwicklung kann als Indikator erfasst werden.

Cité des Marktes

Im Gegensatz zur staatsbürgerlichen Cité kommt dem **Individualismus** in der Cité des Marktes durchaus eine Bedeutung zu. Diese Cité ist besonders durch **Konkurrenz** geprägt, die durch die ökonomische Grundannahme der Begierde nach knappen Ressourcen entsteht und demnach durch Angebot und Nachfrage geprägt wird. Demnach treten zudem **Wettbewerbe und Rivalität** zutage, die Ausdruck des Strebens nach **Reichtum, Besitz und Erfolg** sind (Bogusz, 2010; Boltanski & Thévenot, 1991/2018). In der Praxis äußern sich marktwirtschaftliche Rechtfertigungen daher häufig in einer starken Wettbewerbs- und Kundenorientierung (Nachtwey & Seidl, 2021). Übertragen auf das Bildungssystem soll der Wettbewerb zudem dazu führen, dass der Bildungserfolg verbessert wird (Imdorf, 2011). Dies bedingt zugleich, dass Akteur:innen entweder in die Kategorie der Gewinner:innen oder der Verlierer:innen fallen (Knoll, 2012).

Für die Rechtfertigung des KI-Einsatzes wird insbesondere die internationale Wettbewerbsfähigkeit unterstrichen. Somit sind Vergleiche mit anderen Ländern zu erwarten, aber auch Bestrebungen, den Profit zu erhöhen und sich von anderen Ländern abzuheben, die als Konkurrent:innen betrachtet werden.

Cité der Industrie

Obwohl gewisse Parallelen zwischen der Cité der Industrie und der des Marktes zu liegen scheinen, betonen Boltanski und Thévenot (2018 [1991]) die Andersartigkeit und kritisieren im Zuge dessen gleichsam gegenteilige Auffassungen in der Soziologie (S. 264). Denn nicht das Streben nach Reichtum ist hier die Antriebsfeder des Handelns, sondern die **Innovation, Funktionalität und Effizienz**. Um kontinuierlichen **Fortschritt und Optimierung, Leistung und Produktivität** zu erreichen, bedarf es jedoch auch eines gewissen Maßes an Stabilität und Kontrolle (Bogusz, 2010). In diesem Sinne ist die Cité der Industrie auch durch eine **Metrisierung** in Form von Kennziffern und Kriterien geprägt, die die Formulierung klarer Zielvorgaben gestatten. Daraus entstünde eine gewisse zeitliche Äquivalenz, die sich besonders in Tools widerspiegelt, „die sich wie etwa Programme, Pläne oder Budgets dazu eignen, die Zukunft zu managen", so Boltanski und Thévenot (2018 [1991], S. 282, Herv. i. O.). Übertragen auf das Bildungssystem ist die Cité der Industrie stark outputbezogen und verfolgt **leistungsideologische Ziele**. Demnach strebt sie – im Gegensatz zur staatsbürgerlichen Cité – **explizit nicht nach Chancengleichheit** für alle. Im Gegenteil: „Mit homogenisierten Leistungsklassen will sie verhindern, dass Schüler mit Lernschwierigkeiten in Klassen mit höheren Leistungsanforderungen den Lernrhythmus verlangsamen und die Leistungszuwächse schmälern könnten" (Imdorf, 2011, S. 235).

Wenngleich der **technologische und wissenschaftliche Fortschritt** nicht uneingeschränkt der industriellen Cité vorbehalten ist, wird diese doch traditionell hier verortet (Knoll, 2012). So wird technologischer Fortschritt aus Sicht dieser Cité als unabdingbar gesehen. Technische Hilfsmittel sollen in diesem Sinne dazu dienen, den Bildungssektor effizienter zu gestalten und den erforderlichen **Ressourceneinsatz** zu **minimieren**. Darüber hinaus können sie ebenso dazu genutzt werden, **Planungssicherheit** für einzelne Institutionen zu schaffen, beispielsweise in Form von Prognosen zukünftiger Studierendenzahlen (Imdorf, 2011).

Die projektbasierte Cité

In „Der neue Geist des Kapitalismus" erläutern Boltanski und Chiapello (2003) die Entstehung der projektbasierten Cité, welche maßgeblich durch neoliberale Entwicklungen geprägt sei. Anders als die sechs zuvor beschriebenen Cités begründen die Autoren die projektbasierte Cité nicht unter Bezugnahme auf historische Vertreter:innen der politischen Philosophie, sondern als Reaktion auf soziale Transformationsprozesse im Rahmen der Kapitalismuskritik (Blättler, 2008, S. 81).

Maßgeblich durch die von der 68er-Bewegung vorgebrachte Kritik an Ungleichheit, autoritären Strukturen, starren Hierarchien, Unterdrückung und der „Uniformierung in einer Massengesellschaft“ (Boltanski & Chiapello, 2001, S. 468) geprägt, sei ein *kapitalistischer Geist* entstanden, der eine neue Rechtfertigung als Legitimitätsgrundlage gegenwärtiger Strukturen schaffe (Blättler, 2008; Diaz-Bone, 2022). Dabei „fungiert der kapitalistische Geist [bei Weber] als motivationale und legitimatorische Bedingung der Möglichkeit für die lebenspraktische Durchsetzung einer rationalen Wirtschaftsgesinnung“ (Nachtwey & Seidl, 2021, S. 241). So beschreiben Boltanski und Chiapello (2003) ein ideologisches Konstrukt, welches die Verfolgung eines höheren Sinns beschreibt, der über die Kapitalakkumulation hinausgeht, um das Engagement insbesondere der Arbeitnehmer:innen in einem kapitalistischen System zu erklären, in dem es für einfache Bürgerinnen und Bürger nicht viel zu gewinnen gibt.

Zentral sind dabei der **Abbau starrer Hierarchien** auf der einen Seite und die Entwicklung netzwerkartiger Arbeitsstrukturen auf der anderen Seite (Bogusz, 2010). Im Gegensatz zur häuslichen Cité wird das Tradierte abgelehnt und **Veränderung stellt einen Wert an sich dar**, Stabilität und Verwurzelung werden dagegen als hinderlich angesehen. Insofern wird auch kein übergeordnetes Ziel verfolgt, sondern vielmehr eine Reihe von Projekten, die nach Abschluss durch neue ersetzt werden. Dabei spielen **Mitgestaltung und Teamarbeit** eine Rolle, aber auch **Autonomie, Freiheit und Authentizität** (Boltanski & Chiapello, 2001). Rechtfertigungen aus dem Blickwinkel der projektbasierten Cité rücken also einzelne Probleme und die Erarbeitung ihrer Lösung in den Fokus, ohne dabei nach allgemeiner Gültigkeit zu streben. Rechtfertigungen des Einsatzes von KI sind in diesem Sinne als anwendungsbezogen zu verstehen und nicht darauf bedacht, beispielsweise eine allgemeine „vertrauenswürdige KI“ (Europäische Kommission, 2019) zu garantieren. Stattdessen müssen sie in einer spezifischen Situation dazu dienen, einen spezifischen Zweck zu erreichen. Zur Erreichung dieses Zwecks kann es vonnöten sein, Expertise ‚von außen‘ einzuholen und in diesem Sinne ein neues Netzwerk zu bilden. Ist das Ziel jedoch erfüllt, können neue Projekte, das heißt neue KI-Lösungen für andere Anwendungsfelder, betrachtet werden, die wiederum den Aufbau neuer Netzwerke erforderlich machen können.

Cité der Ökologie

Auch die Cité der Ökologie oder auch grüne Cité ist mit der Reaktion auf globale Umweltbewegungen, gewissermaßen als Folge sozialen Wandels und gesellschaftlicher Kritik, zu betrachten (Lafaye & Thévenot, 1993; Thévenot et al., 2000) und erlangt somit in heutigen Debatten möglicherweise höhere Bedeutung als je zuvor. Analog zu Forderungen derzeitiger Klimabewegungen wie *Fridays for Future* stehen Werte der **Nachhaltigkeit und Erneuerbarkeit** im Vordergrund. Aber auch das Wohlbefinden und die **Gesundheit** werden thematisiert. Der Blick ist dabei nach vorne gerichtet und potenzielle Szenarien und Auswirkungen auf **zukünftige Generationen** werden berücksichtigt (Thévenot et al., 2000). Rechtfertigungen, die die Einzigartigkeit der Natur ignorieren und Konsum oder Verschwendung fördern, werden nicht toleriert (Hildebrandt, 2018). Gleichzeitig wird die Cité der Ökologie jedoch auch mit Technikforschung und Wissenschaft verbunden, indem beispielsweise immer wieder Grenzwerte gesetzt werden oder Treibhausgasemissionen verringert werden sollen (Knoll, 2012, S. 70).

Cité der Solution

Die Cité der Solution greift die von Morozov (2013) kritisierte Auffassung des Solutionismus auf, dass ein **jedes Problem als technisches Problem** und im Besonderen als technisch zu lösendes Problem zu verstehen sei:

> „[S]olutionism presumes rather than investigates the problems that is trying to solve, reaching ‚for the answer before the questions have been fully asked.‘ (…) Solutionism,

> thus, is not just a fancy way of saying that for someone with a hammer, everything looks like a nail; it's not just another riff on the inapplicability of 'technological fixes' to 'wicked problems' (…). It's not only that many problems are not suited to the quick-and-easy solutionist tool kit. It's also that what many solutionists presume to be 'problems' in need of solving are problems at all; a deeper investigation into the very nature of these 'problems' would reveal that the inefficiency, ambiguity, and opacity–whether in politics or everyday life–that the newly empowered geeks and solutionists are rallying against are not in any sense problematic." (Morozov, 2013, S. 6)

Wenngleich Morozov dieses Phänomen nicht als gänzlich neues betrachtet, betont er dennoch den Bedeutungsgewinn des Solutionismus im Zusammenhang mit dem digitalen Wandel und dem Machtgewinn großer Internetkonzerne sowie insbesondere der Entstehung des Silicon Valleys (Morozov, 2013, S. 14). So sei der gegenwärtige Kapitalismus beziehungsweise die Kapitalakkumulation untrennbar mit neuen Technologien verbunden (Morozov, 2016). Nachtwey und Seidl (2021) fassen auf dieser Basis zusammen, dass die Entstehung eines digitalkapitalistischen Geistes zu beobachten sei, die sich gewissermaßen als Weiterentwicklung des kapitalistischen Geistes der projektbasierten Cité konstituiere. Der Geist des digitalen Kapitalismus reagiere „einerseits auf Verschiebungen des Akkumulationsregimes in Form der Herausbildung des digitalen Kapitalismus; andererseits auf Legitimationsprobleme eines moralisch und politisch entfesselten Finanzmarkt-Kapitalismus, wie sie – seit der weltweiten Finanz- und Wirtschaftskrise mit zunehmender Schärfe – vor allem von Seiten der Sozialkritik artikuliert werden" (Nachtwey & Seidl, 2021, S. 242).

Auf Basis dieser Kritik knüpft das solutionistische Verständnis auch an das **technokratische Bewusstsein** an, das auf der einen Seite technisch-wissenschaftliche Expertise in den politischen Entscheidungsraum einbindet – auf der anderen Seite jedoch Entpolitisierungstendenzen aufweist, indem gesellschaftliche Probleme und politisch-demokratische Entscheidungsbefugnisse durch technologische Antworten substituiert werden (Nachtwey & Seidl, 2021). Vor dem Hintergrund dieser geradezu technikverherrlichenden Weltanschauung kann die Cité der Solution in besonderem Maße als Rechtfertigungsnarrativ der Einführung künstlich intelligenter Systeme herangezogen werden. Anknüpfend an Morozov (2013) bedarf es demnach zuerst einer Problemartikulation – unabhängig von der tatsächlichen gesellschaftlichen Bedeutung des Problems – und anschließend der Vorstellung der **technologischen Lösbarkeit**. Dabei wird diese Lösung entweder als einzig mögliche Option oder zumindest als die bestmögliche Option dargestellt und bietet daher aus Sicht der Solutionist:innen die legitimatorische Grundlage der Technikimplementierung. Von besonderer Bedeutung ist dabei die Auffassung, zur Weltverbesserung beizutragen. Dabei ist Stillstand negativ konnotiert und **Veränderung stets positiv.** Institutionen und Regularien gelten gleichzeitig als Hindernis für die Optimierungsbestrebungen, weil sie der Innovation im Wege stehen (Nachtwey & Seidl, 2021).

Literaturverzeichnis

Blättler, S. (2008). Stichwort: Luc Boltanski - Eine Soziologie der Moral. *WestEnd - Neue Zeitschrift für Sozialforschung*, *5*(2), 79–83.

Bogusz, T. (2010). *Zur Aktualität von Luc Boltanski: Einleitung in sein Werk*. VS Verlag für Sozialwissenschaften.

Boltanski, L. & Chiapello, E. (2001). Die Rolle der Kritik in der Dynamik des Kapitalismus und der normative Wandel. *Berliner Journal für Soziologie*, *11*(4), 459–477. https://doi.org/10.1007/BF03204032

Boltanski, L. & Chiapello, È. (2003). *Der neue Geist des Kapitalismus*. *Édition Discours: Bd. 30*. UVK-Verlagsgesellschaft.

Boltanski, L. & Thévenot, L. (2018 [1991]). *Über die Rechtfertigung: Eine Soziologie der kritischen Urteilskraft* (2. Auflage). Hamburger Edition.

Diaz-Bone, R. (2022). Soziologie der Konventionen. In H. Delitz (Hrsg.), *Soziologische Denkweisen aus Frankreich* (S. 471–493). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36949-1_18

Europäische Kommission. (2019). *Ethik-Leitlinien für eine vertrauenswürdige KI.* Hochrangige Expertengruppe für künstliche Intelligenz; Europäische Kommission. https://doi.org/10.2759/22710

Forst, R. (2015). *Normativität und Macht: Zur Analyse sozialer Rechtfertigungsordnungen*. Suhrkamp.

Gadinger, F. & Yildiz, T. (2012). Unruhige Legitimität. Die Finanzkrise aus Sicht der pragmatischen Soziologie der Kritik. In A. Geis, F. Nullmeier & C. Daase (Hrsg.), *Der Aufstieg der Legitimitätspolitik: Rechtfertigung und Kritik politisch-ökonomischer Ordnungen* (S. 302–319). Nomos. https://doi.org/10.5771/9783845245072-302

Hartmann, M. (2008). Rechtfertigungsordnungen und Anerkennungsordnungen. Zum Vergleich zweier Theoriemodelle. *WestEnd - Neue Zeitschrift für Sozialforschung*, *5*(2), 104–119.

Hildebrandt, K. (2018). Der Mediendiskurs zu Umweltthemen: Eine vergleichende Analyse der Berichterstattung über den Emissionshandel in Deutschland und Großbritannien. In M. Lünenborg & S. Sell (Hrsg.), *Politischer Journalismus im Fokus der Journalistik* (S. 191–213). Springer VS.

Imdorf, C. (2011). Rechtfertigungsordnungen der schulischen Selektion. In S. K. Amos, W. Meseth & M. Proske (Hrsg.), *Öffentliche Erziehung revisited* (S. 225–245). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92615-5_11

Knoll, L. (2012). *Über die Rechtfertigung wirtschaftlichen Handelns: CO2-Handel in der kommunalen Energiewirtschaft*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-18983-3

Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung: Grundlagentexte Methoden* (5. Auflage). *Grundlagentexte Methoden*. Beltz Juventa.

Lafaye, C. & Thévenot, L. (1993). Une justification écologique? Conflits dans l'aménagement de la nature. *Revue Française de Sociologie*, *34*(4), 495–524. https://doi.org/10.2307/3321928

Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13., überarbeitete Auflage). Beltz.

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. Public Affairs/Perseus Books.

Morozov, E. (3. Juli 2016). Beware the 'empathy-washing' of self-proclaimed caring capitalists | Evgeny Morozov | The Guardian. *The Guardian*. https://www.theguardian.com/commentisfree/2016/jul/02/beware-technology-giants-claiming-compassion-for-refugees-evgeny-morozov

Nachtwey, O. & Seidl, T. (2021). Ideologie und der Geist des digitalen Kapitalismus. In W. Matiaske & W. Nienhüser (Hrsg.), *Ökonomie und Gesellschaft: Jahrbuch 32. Ökonomie und Ideologie* (S. 235–266). Metropolis-Verlag.

Peter, T. (2015). Die Gerechtigkeit der Spitze. Über die Rechtfertigung von Exellenz im deutschen Bildungsdiskurs. In M. Dammayr, D. Graß & B. Rothmüller (Hrsg.), *Sozialtheorie. Legitimität: Gesellschaftliche, politische und wissenschaftliche Bruchlinien der Rechtfertigung* (S. 177–197). transcript-Verlag.

Schubert, K. & Klein, M. (2020). *Das Politiklexikon.* Lizenzausgabe Bonn: Bundeszentrale für politische Bildung (7., aktual. u. erw. Aufl.). Dietz.

Thévenot, L., Moody, M. & Lafaye, C. (2000). Forms of Valuing Nature: Arguments and Modes of Justification in French and American Environmental Disputes. In M. Lamont & L. Thévenot (Hrsg.), *Cambridge Cultural Social Studies. Rethinking comparative cultural sociology: Repertoires of evaluation in France and the United States* (2. ed., repr, S. 229–272). Cambridge University Press. https://doi.org/10.1017/CBO9780511628108.009

A6. Reliabilitätsprüfung der Codierung der qualitativen Inhaltsanalyse

		Codier:innenpaar			Durchschnittliche Übereinstimmung	
Reliabilitätstest-Nr.	**Dokument**	**A/B**	**A/C**	**B/C**	**Cohen's Kappa r_K**	
Reliabilitätstest I (10.05.2023)	GWK_1052, 2020	0,939	0,880	0,860	0,893	0,893
Reliabilitätstest II (16.05.2023)	GEW_1218, 2017	1,000	0,950	0,950	0,967	0,920
	LT-SN_894, 2021	0,900	0,860	0,880	0,880	
Reliabilitätstest III (26.07.2023)	HFD_1214, 2021	0,960	0,950	0,950	0,953	0,925
	LT-SH_670, 2019	0,950	0,860	0,880	0,897	

Anmerkung: Berücksichtigt wurde die Übereinstimmung der vergegeben Codes innerhalb eines Dokuments, wobei nicht vergebene Codes als Übereinstimmung gewertet wurden.

A7. Relative Häufigkeit der Valenz der Rechtfertigung nach Akteursgruppen

Akteur:in	Pro KI	Abwägend Positiv	Keine Valenz	Abwägend Negativ	Contra KI	n
Politische Akteur:innen						
Die Linke	19,4%	19,4%	9,7%	9,7%	41,9%	31
Die Grünen	59,6%	21,2%	3,8%	11,5%	3,8%	52
SPD	81,3%	12,0%	0,0%	2,7%	4,0%	75
CDU/CSU	83,1%	9,9%	2,8%	2,8%	1,4%	71
FDP	94,5%	3,6%	1,8%	0,0%	0,0%	55
AfD	35,7%	28,6%	0,0%	7,1%	28,6%	28
Andere Partei	80,0%	0,0%	0,0%	0,0%	20,0%	5
Regierungen	77,8%	9,2%	8,7%	1,4%	2,9%	207
Parlamentarische Gremien	63,3%	19,5%	7,8%	2,3%	7,0%	128
Bund-Länder-Zusammenschlüsse	93,1%	3,4%	3,4%	0,0%	0,0%	87
Wissenschaftliche Akteur:innen	59,7%	22,0%	6,6%	2,9%	8,8%	273
Gewerkschaftliche Akteur:innen	2,4%	7,3%	7,3%	19,5%	63,4%	41
Verbände	52,5%	20,8%	10,9%	1,0%	14,9%	101
Wirtschaftliche Akteur:innen	81,8%	0,0%	9,1%	0,0%	9,1%	11

Anmerkung: Dargestellt wird die Verteilung der Valenzen der Rechtfertigungsbezüge innerhalb der Akteursgruppen. Für jede Akteursgruppe dient die jeweilige Zeilensumme als Gesamtzahl der identifizierten Valenzzuordnungen innerhalb dieser Kategorie und bildet die Referenz zur Berechnung der relativen Anteile. | Eine vertiefende Ausdifferenzierung der Akteursgruppen (z. B. wissenschaftliche Akteur:innen oder Bund-Länder-Zusammenschlüsse) kann dem digitalen Anhang B3 entnommen werden.

A8. Relative Häufigkeit der Cité-Bezüge nach Akteursgruppen

Akteur:in	Inspiration	Haus	Meinung	Staatsbürger	Industrie	Markt	Projekt	Ökologie	Solution	Sonstige	Keine Rechtfertigung	n
Politische Akteur:innen												
Die Linke	1,5%	16,4%	3,0%	38,8%	9,0%	9,0%	6,0%	4,5%	7,5%	1,5%	3,0%	67
Die Grünen	0,8%	6,5%	1,6%	30,9%	9,8%	22,8%	11,4%	7,3%	5,7%	0,0%	3,3%	123
SPD	1,4%	8,1%	2,7%	22,3%	16,9%	16,2%	8,1%	2,0%	12,2%	0,0%	10,1%	148
CDU/CSU	2,9%	4,1%	2,3%	15,7%	18,0%	25,0%	7,6%	2,9%	17,4%	0,0%	4,1%	172
FDP	0,6%	4,2%	3,0%	20,8%	20,8%	18,5%	5,4%	1,8%	18,5%	0,0%	6,5%	168
AfD	1,4%	13,7%	1,4%	28,8%	8,2%	27,4%	4,1%	1,4%	12,3%	0,0%	1,4%	73
Andere Partei	0,0%	10,0%	10,0%	10,0%	20,0%	20,0%	10,0%	0,0%	20,0%	0,0%	0,0%	10
Regierungen	0,3%	2,9%	3,9%	22,1%	22,8%	16,6%	4,9%	3,3%	6,5%	0,0%	16,6%	307
Parlamentarische Gremien	1,2%	6,3%	3,9%	32,0%	21,9%	15,6%	6,9%	3,0%	8,4%	0,0%	0,9%	334
Bund-Länder-Zusammenschlüsse	0,0%	0,0%	3,7%	13,9%	33,3%	19,4%	6,5%	0,0%	10,2%	0,9%	12,0%	108
Wissenschaftliche Akteur:innen	0,0%	9,5%	3,8%	19,5%	25,4%	17,1%	9,0%	0,2%	7,4%	1,4%	6,7%	421
Gewerkschaftliche Akteur:innen	3,2%	17,7%	0,0%	50,0%	11,3%	9,7%	4,8%	0,0%	1,6%	1,6%	0,0%	62
Verbände	1,3%	9,4%	4,3%	25,5%	25,1%	12,8%	7,7%	0,4%	8,5%	1,7%	3,4%	235
Wirtschaftliche Akteur:innen	0,0%	0,0%	0,0%	15,4%	23,1%	38,5%	7,7%	0,0%	15,4%	0,0%	0,0%	13

Anmerkung: Dargestellt wird die Verteilung der Cité-Bezüge innerhalb der Akteursgruppen. Für jede Akteursgruppe dient die jeweilige Zeilensumme als Gesamtzahl der identifizierten Cité-Bezüge innerhalb dieser Kategorie und bildet die Referenz zur Berechnung der relativen Anteile. | Eine vertiefende Ausdifferenzierung der Akteursgruppen (z. B. wissenschaftliche Akteur:innen oder Bund-Länder-Zusammenschlüsse) kann dem digitalen Anhang B3 entnommen werden.

Eidesstattliche Versicherung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der ‚Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf' erstellt worden ist.

Düsseldorf, den 18. Dezember 2025 Birte Keller[1]

[1] Abdruck der Unterschrift entfällt aus datenschutzrechtlichen Gründen.