

Aus dem Labor für Molekulare Neurobiologie
Neurologische Klinik
der Heinrich-Heine-Universität Düsseldorf
Leiter: Univ.-Prof. Dr. Hans Werner Müller

Die SCIO-Ontologie:
Eine Formalisierung von Konzepten aus
präklinischen Studien über traumatische
Rückenmarksverletzungen

Dissertation
zur Erlangung des Grades eines Doktors der Zahnmedizin
der Medizinischen Fakultät der Heinrich-Heine-Universität
Düsseldorf

vorgelegt von
Tarek-Fabian Kirchhoffer
2020

Als Inauguraldissertation gedruckt mit Genehmigung der Medizinischen
Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:

Dekan/in: Prof. Dr. med. Nikolaj Klöcker

Erstgutachter/in: Prof. Dr. rer. nat. Hans Werner Müller

Zweitgutachter/in: Prof. Dr. rer. nat. Stefan Conrad

Teile dieser Arbeit wurden veröffentlicht:

Paassen, Benjamin; Stöckel, Andreas; Dickfelder, Raphael; Göpfert, Jan Philip; Brazda, Nicole; Kirchhoffer, Tarek; Müller, Hans Werner; Klinger, Roman; Hartung, Matthias; Cimiano, Philipp, 2014, Ontology-based Extraction of Structured Information from Publications on Preclinical Experiments for Spinal Cord Injury Treatments, Third Workshop on Semantic Web and Information Extraction (SWAIE), The 25th International Conference on Computational Linguistics (COLING), Association for Computational Linguistics and Dublin City University, Dublin, Ireland

I Zusammenfassung

Das aus präklinischen Tierstudien über traumatische Rückenmarksverletzungen (TRV) in hoher Quantität generierte Wissen liegt in natürlicher Sprache vor und ist somit nicht für automatische Metastudien zugänglich und anwendbar. Die Entwicklung geeigneter Therapien in der Domäne des Neurotraumas wird nachhaltig und negativ durch das fehlende Bewusstsein in Wissenschaftskreisen über das bereits vorhandene Wissen, das aus einer Vielzahl an präklinischen Studien hervorgeht, beeinflusst. Zur Bewältigung dieses Problems wurde die Spinal Cord Injury Ontology (SCIO) entwickelt, um das bereits vorhandene Wissen über therapeutische Konzepte und insbesondere die Beschreibung relevanter Parameter und Ergebnisse aus der großen Menge an präklinischer Literatur einzufangen. SCIOs kurzfristiges Ziel ist die Bereitstellung einer zielorientierten Datenextraktion, um eine präklinische Studie über traumatische Rückenmarksverletzungen vollständig und detailliert zu repräsentieren. Langfristig soll die SCIO als Grundlage für die automatische Datenextraktion aus wissenschaftlichen Publikationen dienen.

Aus 140 Publikationen (*peer-reviewed*) wurden Konzepte, wie die verwendeten Untersuchungsmethoden, Tierspezies, Verletzungsmodelle und Therapien manuell extrahiert. Es wurde ein Klassengefüge mit Attributen und einer Taxonomie definiert. Jede Klasse wurde mit dem zentralen Konzept der Ontologie, dem Ergebnis (*Outcome*) eines Experimentes, verknüpft. Basierend auf einer Untersuchungsmethode (z.B. die Überprüfung der Fortbewegungsfähigkeit) werden unterschiedliche experimentelle Gruppen in einem direkten Vergleich durch die Ontologie als „Veränderung“ (*change*) oder „keine Veränderung“ (*no change*) eingestuft. Bei der Datenextraktion mit SCIO werden die beschriebenen Ergebnisse einer wissenschaftlichen Publikation mit (1) den gegenseitig verglichenen experimentellen Gruppen verknüpft, die wiederum Auskunft über das experimentelle Setup und (2) die angewendeten Untersuchungsmethoden geben. Die manuelle Annotation von wissenschaftlichen Texten verlief simultan zur Konstruktion der Ontologie, um ein notwendiges Maß an Konzepten für eine adäquate Repräsentation der Domäne durch SCIO zu gewährleisten. Als Proof of Concept wurde die SCIO durch zwei Diplom-Biologen für die semi-automatische Datenextraktion auf 22 wissenschaftliche Publikationen angewendet.

Die SCIO wurde konzipiert, um alle Hauptparameter und Ergebnisse einer präklinischen Studie über TRV zu extrahieren. Die Ontologie besteht aus 385 Klassen und 46 Attributen. Die Granularität der Ontologie entspricht den Anforderungen die Domäne der TRV übersichtlich und detailliert zu repräsentieren. In einem mit Diplom-Biologen durchgeführten Annotationsprojekt validierten hohe Übereinstimmungen bei der semi-automatischen Datenextraktion mit SCIO ihre Effizienz. Darüber hinaus wurde die SCIO von Paasen et al. (2014) als Grundlage für die erfolgreiche automatische Datenextraktion aus wissenschaftlichen Publikationen über TRV verwendet. Im weiteren Verlauf soll die Ontologie den Translationsprozess von erfolgsversprechenden präklinischen Studien in klinische Studien fördern.

II Abstract

A vast amount of knowledge generated by preclinical research in spinal cord injury (SCI) is described in natural-language publications and is thus not accessible to automated meta-studies. The design of therapeutic concepts in the neurotrauma domain is notoriously affected by the lack of complete knowledge about the large amount of data produced by experiments in preclinical studies. To overcome this problem the Spinal Cord Injury Ontology (SCIO) was developed in order to support the systematic capture of knowledge about therapeutic concepts, particularly the description of relevant parameters and results of experiments in the vast preclinical literature available on the topic. SCIO's short-term goal is to supply scientists with a targeted form of data-extraction in order to represent a preclinical study in its entirety for a more transparent and detailed overview of the experimental setup and results. In the long-run SCIO serves as a blueprint for data mining, a method for automated data extraction.

Descriptions and concepts from more than 140 peer-reviewed publications on spinal cord injury in animal models were manually extracted, e.g., types of investigation methods, animal strains or experimental lesion types. Classes, relations and a taxonomy were defined and each class was linked to the central concept of the ontology, i.e., the result of one experiment in a study, which is formalized as a "change" or "no change" of a certain parameter (e.g., locomotor function) between two experimental conditions (e.g. treated versus untreated). In SCIO, the results described in one publication are linked (1) to their specific experimental groups containing information about the experimental setup and (2) the investigation method that was used to evaluate the experiment's result. Construction of the ontology was accompanied by textual annotations to cover a substantial amount of natural language terms that are used in scientific publications in order to yield a reference to the concepts contained in SCIO. Finally, the ontology was applied in terms of a proof of concept by two biologists who annotated 22 scientific publications about preclinical spinal cord injury experiments.

SCIO has been designed to capture the main parameters and results of preclinical studies in the spinal cord injury domain. The final ontology consists of 385 classes and 46 relations. Based on the ongoing scientific discussion, the data granularity suits the actual domain needs for objective level-of-evidence estimation on therapeutic effects. SCIO represents all relevant aspects of a preclinical study about spinal cord injury and provides a straightforward approach towards information extraction. The annotation project carried out with two biologists yielded a high level of agreement amongst them in terms of data extraction with SCIO which proved its successful application. Furthermore, SCIO served as a template for an ontology-based automated data extraction from preclinical literature as proven by Paasen et al. (2014) who was able to use the ontology's structure for datamining. Thus, the SCIO can serve as a tool to increase transparency and exchange of data amongst scientists in order to expedite the transformation of preclinical studies into clinical studies to yield success promising therapy options for patients who suffer from traumatic spinal cord injuries.

III Abkürzungen

CDE	Common Data Elements
CITEC	Cluster of Excellence Cognitive Interaction Technology
CNR	Center for Neuronal Regeneration
ECNP	European College of Neuropsychopharmacology
IRR	Interrater reliability
NIH	National Institute of Health
NINDS	National Institute of Neurological Disorders and Stroke
NIST	National Institute of Standards and Technology
OAVT	Object Attribute Value Triplet
SCI	Spinal Cord Injury
SCIO	Spinal Cord Injury Ontology
TRV	traumatische Rückenmarksverletzungen
UML	Unified Modeling Language

IV Inhaltsverzeichnis

1 Einleitung	1
1.1 Problemstellung	2
1.2 Wissenschaftlicher Hintergrund	2
1.2.1 Die fehlende Transparenz	3
1.2.2 Der fehlende Informationsaustausch	5
1.2.3 Die erschwerte Reproduzierbarkeit	6
1.3 Der aktuelle Forschungsstand	9
1.4 Ziele der Arbeit	11
2 Material und Methoden	14
2.1 Definition des Kompetenzbereiches	15
2.2 Konstruktion der SCIO	18
2.2.1 Datenextraktion aus wissenschaftlicher Literatur	19
2.2.2 Konzeptualisierung und Formalisierung der Domäne	20
2.3 Der Aufbau der Spinal Cord Injury Ontology	31
2.3.1 Das Konzept der Klassenhierarchien	31
2.3.2 Die Hauptklasse Animal	33
2.3.3 Die Hauptklasse Injury	34
2.3.4 Die Hauptklasse Treatment	36
2.3.5 Die komplementären Klassen	38
2.3.6 Die Hauptklasse Investigation Method	41
2.3.7 Die Hauptklasse Outcome	44
2.4 Attribute und das Relationsgefüge der SCIO	45
2.5 Implementierung und Evaluation der Ontologie	50
2.5.1 Die Anwendungsrichtlinien	50
2.5.2 Das Annotationsprojekt	60

3 Ergebnisse	62
3.1 Die relationale Struktur und Anwendung der SCIO	62
3.2. Analyse und Evaluation des Annotationsprojektes	63
3.2.1 Der Interrater reliability-Wert (P_0)	64
3.2.2 Der Cohen´s Kappa Koeffizient (κ)	65
4 Diskussion	68
4.1 Erkenntnisse aus dem Annotationsprojekt	68
4.2 Erweiterbare Anwendungsbereiche der SCIO	70
4.3 Schlussfolgerungen	71
4.4 Einfluss der SCIO auf die gute wissenschaftliche Praxis	73

Einleitung

1.1 Problemstellung

Verletzungen des zentralen Nervensystems (ZNS) resultieren in einem permanenten Verlust von motorischen, sensorischen und autonomen Funktionen, da im Vergleich zu anderen Organen die regenerative Kapazität des ZNS sehr eingeschränkt ist [30]. Patienten bei Zustand nach vorausgegangenen Schlaganfällen, traumatischen Hirn-, oder Rückenmarksverletzungen leiden oft ein Leben lang unter körperlichen Einschränkungen, die neben der Auslösung neuropathischer Schmerzen einen Einfluss auf grundlegende Körperfunktionen, Sinnesempfindungen, Bewegung und Sprache haben. Neben der eingeschränkten Teilnahme am gesellschaftlichen Leben, verursacht durch die physischen und psychischen Begleiterscheinungen einer traumatischen Rückenmarksverletzung (TRV), spielen sozioökonomische Faktoren, wie die finanzielle Belastung des Gesundheitssystems, eine Rolle [9]. Zurzeit beläuft sich die Therapie traumatischer Rückenmarksverletzungen lediglich auf die Minimierung der bestehenden Beschwerdesymptomatik und Aufrechterhaltung der noch vorhandenen körperlichen Funktionen [28]. Trotz der äußerst beschwerlichen Lebensumstände von Patienten bei Zustand nach TRV, stehen keine adäquaten Therapien für eine Rehabilitation in Aussicht. Dies ist darauf zurückzuführen, dass im Verlauf der letzten Jahrzehnte die Diskrepanz zwischen den präklinischen bzw. tierexperimentellen Forschungsbemühungen und dem eigentlichen Voranschreiten bei der Identifikation effektiver Therapien von TRV zugenommen hat [13]. Trotz der hohen Anzahl an durchgeführten präklinischen Studien konnten in den vergangenen Jahrzehnten keine entwicklungsfähigen Ergebnisse produziert werden, um eine Translation von erfolgsversprechenden präklinischen Experimenten in klinische Studien zu gewährleisten. Dieser Aufwand spiegelt sich in der jährlich exponentiell steigenden Anzahl von wissenschaftlichen Veröffentlichungen wider. Heute sind über 150.000 Publikationen zum Thema ZNS-Trauma auf der wissenschaftlichen Plattform PubMed vorhanden. Zum Thema traumatische Rückenmarksverletzungen liegen über 40.000 Publikationen vor, die einen enormen Korpus an präklinischen Tierstudien beinhalten. Das aus diesen wissenschaftlichen Veröffentlichungen hervorgehende Fachwissen liegt lediglich in unstrukturierter Form, respektive in natürlicher Sprache vor, und ist somit nicht für automatische Metastudien, z.B. im Sinne einer Datenbank, zugänglich.

Darüber hinaus ist bei der bestehenden Informationsflut, aktuell durchschnittlich acht peer-referierte Publikationen täglich (PubMed), von einer fehlenden Kenntnisnahme aller relevanten Publikationen durch die Wissensgemeinschaft auszugehen. Hieraus ergibt sich die Problematik, dass es für individuelle Wissenschaftler nahezu unmöglich ist sachkundige, auf dem aktuellen Wissensstand basierende Entscheidungen zu treffen, um weitere evidenzbasierte präklinische Studien durchzuführen, die für eine erfolgreiche Weiterführung in der klinischen Umgebung von Bedeutung sein können. Potenziell birgt dieser Umstand Insuffizienzen bei der Auswahl von experimentellen Strategien, z.B. bei der Auswahl von Tiermodellen, Behandlungskonzepten und bei der Umsetzung von Testprotokollen, was weitreichende Folgen für die zeitgerechte Entwicklung von Therapieoptionen für TRV hat. Angesichts der Schwere der Problematik sind die Ursachen jedoch salient und identifizierbar.

1.2 Der wissenschaftliche Hintergrund

Das saliente Ungleichgewicht zwischen der großen Anzahl an präklinischen Studien und der geringen Ausbeute an vermeintlich entwicklungsfähigen Ergebnissen für klinische Studien ist die Folge einer fehlenden Wahrnehmung von erfolgsversprechenden Therapieansätzen in Wissenschaftskreisen. Die per annum steigende Anzahl an Publikationen verursacht ein nahezu reziprokes Verhältnis zwischen dem stetig zunehmenden Forschungsaufwand und der rückläufigen Identifikation und Weiterentwicklung geeigneter Therapien. Diese Reziprozität klafft kontinuierlich zu Ungunsten einer zeitgerechten Therapieentwicklung auseinander. Dieser Zustand ist sicherlich auch im Paradigma anderer Wissenschaften festzustellen und lässt Rückschlüsse auf den Umgang mit Daten innerhalb von Wissenschaftskreisen ziehen. Hier spielen die Datenverarbeitung, die Präsentation von Daten und deren Bereitstellung eine wichtige Rolle. Im Speziellen lassen sich in der präklinischen Forschung zu TRV die Ursachen für die ausbleibende Fortführung und Weiterentwicklung von wertvollen Therapieansätzen auf den Umgang mit wissenschaftlichen Daten zurückführen. In diesem Zusammenhang kann konstatiert werden, dass studienrelevante Daten und Ergebnisse in einem Format präsentiert werden, das für Wissenschaftler opak, unübersichtlich und nicht beurteilbar ist. Des Weiteren können diese Daten aufgrund eines fehlenden adäquaten Wissensaustausches

nicht für aufbauende Studien zur Bestimmung eines Evidenzgrades verwendet werden. Die unterschiedlichen Ursachen für die Verzögerung bei der Therapiefindung sind in ihrer Qualität einzigartig und gleichzeitig miteinander verknüpft. Die fehlende Transparenz, der fehlende Informationsaustausch und die fehlende Reproduzierbarkeit von bereits durchgeführten experimentellen Studien sind die Hauptursachen für die karge klinische Ausbeute bei der Therapiefindung. Hierbei sind die Übergänge zwischen den Ursachen fließend und durch gegenseitige Wechselwirkungen geprägt.

1.2.1 Die fehlende Transparenz

Das unübersichtliche und zugleich überdurchschnittlich große Volumen an experimentellen Daten, in Form von schriftlich publizierten Methoden, Therapien und Ergebnissen, ist einer der Hauptgründe für die fehlende Transparenz in der präklinischen Forschung über TRV. Eine fehlende Transparenz zeigt sich auch bei der lückenhaften Beschreibung des experimentellen Aufbaus und Ablaufs in wissenschaftlichen Publikationen. Hieraus resultieren Unzulänglichkeiten beim Versuch erfolgversprechende präklinischen Studien zu replizieren. Die Übertragung von präklinischer neurologischer Grundlagenforschung in klinische Studien schreitet unter anderem sehr langsam voran, weil eine große Quantität an unstrukturierten, unübersichtlichen Datensätzen (*Big Data*) unbearbeitet vorliegt [23]. Ein für Wissenschaftler freier und zielgerichteter Zugriff auf diese enormen Datenmengen ist nicht gegeben. Das amerikanische National Institute of Standards and Technology (NIST) definiert mit dem Begriff *Big Data* ein großes Volumen an Datensätzen, das in umfangreicher Vielfalt vorhanden ist, sich mit einer hohen Geschwindigkeit vermehrt, durch ein hohes Maß an Variabilität gekennzeichnet ist und eine skalierbare Datenarchitektur benötigt, um effizient gespeichert, bearbeitet und analysiert zu werden [24]. All dies sind Attribute, die von Wissenschaftlern in der präklinischen Forschung über TRV bei der Publikation von wissenschaftlichen Studien beachtet werden sollten, um den Zugriff auf ein transparentes Wissen zu garantieren. Durch eine Myriade von Möglichkeiten bei der Veröffentlichung von wissenschaftlichen Arbeiten gewinnt *Big Data* in vielen unterschiedlichen Wissenschaften an Bedeutung. Die steigende Anzahl der Internetplattformen und Verlage beschleunigt den zeitlichen Ablauf von der Konzeption einer präklinischen Studie bis zur Veröffentlichung. In diesem Zusammenhang stellt die NIST fest, dass sich im Rahmen der stetig größer werdenden

Datenmengen ein signifikanter, progredienter Anstieg an unstrukturierten Daten zeigt, die ohne jeglichen Bezug zueinander vorliegen und kein Gesamtkonstrukt ergeben [24]. In Kombination mit dem großen Interesse innerhalb der präklinischen Forschergemeinschaft geeignete Therapielösungen zu finden, lässt sich das jährliche exponentielle Wachstum von Publikationen und die damit immer opaker erscheinende Transparenz erklären. Wie Landis et al. (2012) postulierte, ist die Bereitstellung und freie Zugänglichkeit von Wissen der Wegbereiter für den wissenschaftlichen Fortschritt [18]. Beides dient sowohl dem Zweck der Reproduzierbarkeit als auch der Weiterentwicklung von präklinischen Studien. Folglich wird auf diesem Wege die Identifikation robuster Therapieansätze erleichtert. Neben dem hohen Aufkommen an neuen Publikationen und der damit verbundenen Informationsflut suggeriert die Prävalenz unklarer Beschreibungen beim experimentellen Aufbau und bei Versuchsprotokollen in wissenschaftlichen Veröffentlichungen eine notwendige Verbesserung der Berichterstattung [18]. Dies hat weitreichende Auswirkungen auf die Identifikation und Weiterführung von erfolgversprechenden präklinischen Studien.

Zusammengefasst sind die Ursachen für die fehlende Transparenz des bereits vorhandenen Wissens in der Domäne der TRV (1) die hohe Anzahl an individuellen und vom Rest der Wissensgemeinschaft isolierten Experimente, (2) der heterogene experimentelle Aufbau von präklinischen Studien, (3) die individuelle, unabhängige Auswertung von Daten einzelner Labors, (4) die ausbleibende Kenntnisnahme und Verarbeitung der stetig größer werdenden Datensätze unterschiedlicher Herkunft durch die Wissensgemeinschaft und (5) die Entstehung eines eher deskriptiven als unmittelbar verständlichen Wissens [29]. All dies sind Aspekte, die einen strukturierten und zielorientierten intradisziplinären Wissensaustausch innerhalb der Wissensgemeinschaft verhindern. Einige der oben genannten Ursachen, wie der heterogene experimentelle Aufbau oder das schier hohe Aufkommen an präklinischen Studien sind zwangsläufig unvermeidbar und spiegeln die Motivation in der präklinischen Forschergemeinschaft wider, effiziente Therapien herbei zu führen. Dies ist eine durchaus positive Entwicklung, die durch eine strukturierte Präsentation und Bereitstellung von neuen Erkenntnissen zu einer zeitgerechten Umsetzung von Therapieanwendungen führen kann. Hingegen sind die unabhängige Auswertung von Daten, die fehlende Kenntnisnahme und die Präsentation deskriptiver, uneinheitlicher Wissens Elemente Ursachen, die zugunsten des Zugriffs auf ein transparentes Wissen vermieden werden

sollten. Auch hier kann durch eine adäquate Präsentation des bereits vorhandenen Domänenwissens eine klare Übersicht und Erweiterung des aktuellen Wissensstands gewährleistet werden.

1.2.2 Der fehlende Informationsaustausch

Der fehlende Informationsaustausch innerhalb der Wissensgemeinschaft ist die logische Konsequenz der intransparenten Darstellung des aus zahlreichen präklinischen Studien gewonnenen Wissens. Zusätzlich zeigen sich über die Forschergemeinschaft hinaus weitere in ihrem Handeln beeinträchtigte Interessengruppen, wie die pharmazeutische Industrie, Verlage, Journale, wissenschaftliche Plattformen (PubMed), gemeinnützige Organisationen und staatliche Einrichtungen im Gesundheitswesen, die von einem Informationsaustausch, und einer sich daraus entwickelnden evidenzbasierten Studienlage, profitieren würden [31]. Allen voran stehen die durch traumatische Rückenmarksverletzungen betroffenen Patienten, denen keine geeigneten Therapien zu Linderung ihrer Beschwerdesymptomatik in Aussicht gestellt werden. Um die Ursachen für den fehlenden Informationsaustausch innerhalb der Wissenschaftlergemeinschaft zu identifizieren, müssen die vorhandenen Strukturen in der Forschung beleuchtet werden.

In der biomedizinischen Forschung besteht im Vergleich zu anderen Wissenschaften die Aufteilung in einen präklinischen und klinischen Wissenschaftsbereich. Hieraus entstehen besondere Herausforderungen beim Wissensaustausch. Das Wissen muss in diesem Zusammenhang auf zwei Ebenen angemessen weitergeleitet werden. Die erste Ebene entspricht der Kommunikation zwischen Neurowissenschaftlern innerhalb des präklinischen Kompetenzbereiches. Die zweite Ebene bezieht sich auf den Wissensaustausch zwischen Wissenschaftlern aus dem präklinischen und klinischen Umfeld. Während im präklinischen Bereich Naturwissenschaftler wie Biologen und Physiologen den experimentellen Ablauf gestalten, sind es in der Klinik Fachärzte für Neurologie, Neurochirurgie und Orthopädie. Daher ist es nachvollziehbar, dass viele klinische Wissenschaftler nicht mit dem aktuellen experimentellen Aufbau präklinischer Tierstudien und den dazugehörigen Protokollen vertraut sind [16]. Es kann angenommen werden, dass von Seiten der klinischen Wissenschaftler eine adäquate Beurteilung von erfolversprechenden präklinischen Studien erschwert ist und eventuelle Empfehlungen oder Fragen zwischen den unterschiedlichen Kompetenzbereichen erst gar nicht zu Stande kommen. Unter Anbetracht der Tatsache,

dass sich die Fachbereiche nicht adäquat ergänzen, erscheint die Therapieentwicklung weitgehend vorbelastet. Ferner erfordert die Durchführung von klinischen Studien einen hohen finanziellen Aufwand, der sich unter Umständen nicht amortisiert, da die zur Verfügung stehenden, begrenzten finanziellen Mittel nicht gezielt eingesetzt werden und vermeintlich erfolgversprechendere Therapieansätze nicht mehr durchgeführt werden können [16]. Die Bestimmung eines angemessenen Evidenzgrades sollte deshalb eine Voraussetzung für die Fortführung von präklinischen Studien in der klinischen Umgebung sein, um bereits im Vorfeld erfolgversprechende Therapieansätze auswählen zu können. Hierfür ist die Umwandlung des opaken intransparenten Wissens in eine transparente und für Wissenschaftler zugänglichere Form notwendig. Dabei empfiehlt sich im Speziellen die Formalisierung des bereits vorhandenen Wissens durch vordefinierte Konzepte, denen das aus Studien generierte Wissen zugeordnet, und gleichzeitig mit anderen Konzepten aus der Domäne der traumatischen Rückenmarksverletzungen verknüpft werden kann. Diese konzeptualisierte Vernetzung von semantischen Entitäten kann durch die Erstellung einer domänenspezifischen Ontologie gewährleistet werden [24]. Demnach kann das locker geflochtene, deskriptive Wissen in ein zusammenhängendes, und zugleich kompaktes Wissen transformiert werden, das für Wissenschaftler übersichtlicher, verständlicher und zugänglicher ist.

1.2.3 Die erschwerte Reproduzierbarkeit

Die erschwerte Reproduzierbarkeit von präklinischen Studien ist eine aus der fehlenden Transparenz und dem fehlenden Informationsaustausch resultierende Konsequenz. Die ergebnisorientierte Reproduzierbarkeit einer Studie beschreibt die unabhängige Replikation einer bereits durchgeführten Studie unter Anwendung des gleichen experimentellen Aufbaus mit dem Ziel die gleichen Ergebnisse zu produzieren [15]. Analysen haben gezeigt, dass in den USA mehr als 28 Milliarden Dollar jährlich in präklinische Studien investiert werden, von denen sich mehr als 50% nicht replizieren lassen [11]. Eine von der *Nature* Fachzeitschrift mit 1576 Wissenschaftlern durchgeführte Online-Umfrage ergab, dass mehr als 70% der Befragten nicht in der Lage waren die Experimente anderer Wissenschaftler zu replizieren und dass mehr als die Hälfte der Befragten ihre eigenen Experimente nicht replizieren konnten [2]. Im Rahmen der gleichen Umfrage antworteten mehr als 90% der Befragten, dass ein besseres *experimental design*, eine bessere Statistik und eine bessere Betreuung von

Wissenschaftlern die erfolgreiche Replikation von Experimenten unterstützen würde [2].

Die erschwerte Reproduzierbarkeit ist in vielen Wissenschaften ein weit verbreitetes Problem, dessen Prävalenz im Paradigma der präklinischen Forschung höher ist, da es hier im Vergleich zu klinischen Studien keine strengen Regularien und externe Aufsichten gibt, die einen durchsichtigen experimentellen Aufbau und einwandfreien Ablauf fordern [8]. Auch wenn die Herangehensweisen in der präklinischen und klinischen Forschung nicht direkt miteinander vergleichbar sind, was darauf zurückzuführen ist, dass ein gewisses Maß an Kreativität und Entscheidungsfreiheit erforderlich ist, um neue Therapieansätze zu entwickeln, ist zu beachten, dass die Weiterentwicklung des bereits vorhandenen Wissens ein wichtiger Pfeiler des evidenzbasierten Forschungsprinzips ist und für eine erfolgreiche Replikation von präklinischen Studien und somit für die Entwicklung geeigneter Therapien unabdingbar ist [26]. Friedman et al. (2015) sieht die Hauptursachen für nicht reproduzierbare präklinische Studien (1) bei einem inadäquaten experimentellen Aufbau, (2) bei der Verwendung ungeeigneter biologischer Reagenzien und Referenzmaterialien, (3) bei lückenhaften Laborprotokollen und (4) bei einer ungenauen Datenanalyse und Berichterstattung [11]. Eine durch die Autoren von wissenschaftlichen Veröffentlichungen voreingenommene Berichterstattung, gepaart mit unzureichenden Statistiken und wenig Details zum verwendeten Material und den angewandten Methoden, werden ebenfalls als Ursachen für die schlechte Reproduzierbarkeit erwähnt [31]. Weitere einflussnehmende Faktoren auf die erschwerte Replikation von präklinischen Experimenten sind individuelle Unzulänglichkeiten oder fehlende Expertise beim experimentellen Aufbau und fehlende Angaben zu den technischen Details des experimentellen Aufbaus und Ablaufes in wissenschaftlichen Veröffentlichungen [8].

Über 95 % eines befragten Expertenkonsortiums, bestehend aus 250 Wissenschaftlern, sah es für notwendig, dass erfolgversprechende Therapien, die von einer Arbeitsgruppe durchgeführt wurden, zur Verifizierung erneut unabhängig von einer anderen Arbeitsgruppe repliziert werden sollten [16]. Die Replikation von präklinischen Experimenten dient zum einen ihrer Validierung und zum anderen der Weiterentwicklung von erfolgversprechenden Therapieansätzen. Hier ist festzustellen, ob ein Therapieansatz ein gewisses Maß an Robustheit aufweist, um ggf. durch andere

Arbeitsgruppen weiterentwickelt oder bereits in der klinischen Umgebung umgesetzt zu werden. Detaillierte und verständliche Beschreibungen des experimentellen Aufbaus und des Durchführungsprotokolls sind somit Grundvoraussetzungen für eine erfolgreiche Replikation. Diesbezüglich soll bei der experimentellen Berichterstattung die freie Formulierung im Sinne eines freien wissenschaftlichen Geistes aufrechterhalten werden, jedoch ist es notwendig, dass bestimmte Konzepte, die für das allgemeine Verständnis und für die Aussagekraft einer Studie von Bedeutung sind, erwähnt werden sollten. Demnach sollte ein vorgeschriebenes Maß an zu erwähnenden Fachbegriffen als Grundinformation auf wissenschaftlichen Plattformen vorausgesetzt werden. Hierdurch sollten Arbeitsgruppen aufgefordert werden ein Mindestmaß an Evidenz in Ihrer Publikation zu gewährleisten, dass die Chancen auf eine erfolgreiche Replikation und Weiterentwicklung maximiert. Lemmon et al. (2014) schlägt vor, dass diese Grundvoraussetzungen bereits bei der Vergabe von Fördermitteln und bei der Einreichung von Manuskripten für die Publikation von wissenschaftlichen Arbeiten eine Rolle spielen sollten [19].

Bei wissenschaftlichen Publikationen zum Thema TRV wird durch die fehlende Transparenz eine genaue Einsicht in den experimentellen Aufbau einer präklinischen Studie unterbunden und somit eine adäquate Reproduktion derselben erschwert [18]. Dieses Phänomen ist bei einem Großteil der wissenschaftlichen Publikationen über präklinische TRV feststellbar und lässt Rückschlüsse auf die lückenhafte Darstellung der beschriebenen Methoden ziehen. Pharmazeutischen Unternehmen fällt die erschwerte Replikation von präklinischen Experimenten unterschiedlicher Fachbereiche bei Inhouse-Studien auf, so dass die Verlässlichkeit von vermeintlich erfolversprechenden publizierten Daten regelmäßig in Frage gestellt werden muss [27]. Diese Form der Validierung entspricht einem erheblichen Kostenaufwand, der bei erfolgloser Replikation einen wirtschaftlichen Verlust und eine Verzögerung bei der Identifikation geeigneter Therapien verursachen kann. Hierbei konnte in einem unabhängigen Replikationsversuch von 67 publizierten Studien festgestellt werden, dass lediglich 20 bis 25 % der replizierten Daten mit den Ursprungsdaten übereinstimmten [27]. Die geringe Präzision bei der Beschreibung des experimentellen Aufbaus und des Durchführungsprotokolls verursacht mitunter das langsame Voranschreiten bei der Identifikation geeigneter Therapien [11]. Selbst für *Reviewer* ist es bei der Evaluation eingereicherter Manuskripte schwierig, defizitäre Korrelationen zwischen den

beschriebenen Methoden und den präsentierten Ergebnissen festzustellen [18]. Die Identifikation von Fehlerquellen ist somit erschwert, so dass ein Lernprozess basierend auf der Evidenz vorausgegangener Experimente ausbleibt.

Das hohe Aufkommen an Unregelmäßigkeiten beim *study design* und bei der Validierung der verwendeten Forschungswerkzeuge in präklinischen Studien deuten eine gewisse Befangenheit der involvierten Wissenschaftler an, die einer neutralen Berichterstattung entgegenwirkt [26]. Diese Befangenheit ist darauf zurückzuführen, dass Wissenschaftler die Relevanz Ihrer Veröffentlichungen durch die Präsentation von möglichst signifikanten Ergebnissen unterstützen möchten, um somit Ihre Karrierechancen und die finanzielle Unterstützung Ihrer Projekte im weiteren Verlauf zu stärken [26]. Ein durchaus sinnvolles Konzept zur Gewährleistung einer adäquaten Replikation von präklinischen Studien, ist der Vorschlag noch vor Beginn einer Studie Hypothesen und geplante Datenanalysen als Leitfaden schriftlich festzulegen und bei einer unabhängigen Institution einzureichen, um im weiteren Verlauf eine subjektive Auswahl an statistisch signifikanten Ergebnissen zu vermeiden [2]. Praktiken dieser Art sind zur Zeit noch nicht etabliert, so dass der Fokus zur Gewährleistung einer adäquaten Replikation von präklinischen Studien auf das bereits veröffentlichte Wissen gelegt werden sollte, um selektive Verfahren für die Weiterführung von präklinischen Studien zu ermöglichen. Auch hier bietet sich die Präsentation des vorhandenen Domänenwissens über präklinische Studien zum Thema TRV in Form einer Ontologie an, so dass Rückschlüsse auf die verwendeten Materialien und die angewandten Methoden gezogen werden können. Eine konzeptualisierte Zusammenfassung aller Entitäten einer präklinischen Studie über TRV kann bei der Selektion von zu replizierenden Studien hilfreich sein.

1.3 Der aktuelle Forschungsstand

Die eingeschränkte Reproduzierbarkeit und der fehlende Datenaustausch von präklinischen Studien über TRV war der Anlass für Lemon et al. (2014) ein Format zur Erfassung von Details aus präklinischen Tierstudien vorzuschlagen. Das MIASCI - Projekt (*Minimum Information About a Spinal Cord Injury Experiment*) soll mithilfe von 250 definierten Datenelementen relevante Konzepte der präklinischen Forschung über TRV adäquat repräsentieren und in wissenschaftlichen Publikationen für eine genaue Berichterstattung Anwendung finden [20]. Die von derselben Arbeitsgruppe

vorgestellte RegenBase Ontologie verbindet die aus präklinischen Studien gewonnenen Informationen mit Informationen über spezifische biologische Prozesse auf molekularer Ebene, indem sie ihre Konzepte mit den Konzepten anderer Ontologien verbindet [19]. Der hier betriebene Aufwand übersteigt die realistische Zielsetzung das vorhandene Wissen von präklinischen Studien über TRV für Wissenschaftler transparent und zugänglich zu machen. Der Aufwand, der mit der RegenBase Ontologie bei der Repräsentation von molekular biologischen Prozessen betrieben wird, ist bei der Vielzahl von jährlich veröffentlichten Studien nicht umsetzbar. In diesem Zusammenhang wäre die Präsentation einer präklinischen Studie in ihrer Gesamtheit sinnvoller, damit sich Interessengruppen einen Überblick über den experimentellen Aufbau und Ablauf verschaffen können. Kwon et al. (2011) präsentierte ein Punktesystem zur objektiven Bewertung von Publikationen über TRV in präklinischen Tierexperimenten [16]. Eine andere Herangehensweise zur Beurteilung der Effektivität einer Therapie ist das von Kwon et al. (2011) vorgestellte Bewertungsschema [16]: Hier definierte ein Expertenkonsortium aus dem präklinischen Forschungsbereich über TRV spezifische Kriterien zur Beurteilung und Selektion von erfolgversprechenden Therapien aus einer Sammlung von multiplen Therapiestrategien, die für eine Weiterführung in der klinischen Forschung prädestiniert sind. Bei diesen Kriterien spielen zum einen die unabhängige Validierung, im Sinne von erfolgreichen experimentellen Replikationen, und zum anderen die Applikation von Therapien in unterschiedlichen Tier- und Therapiemodellen eine wichtige Rolle. Das Bewertungsschema wurde auf den Volltext von über 100 Publikationen angewendet. Die extrahierten Daten aus den Zusammenfassungen zeigten sich trotz ihrer grobkörnigen Granularität für das Bewertungsschema insuffizient. Wie von Sejnowski et al. (2014) geschildert, hat aufgrund des hohen Aufkommens von "Big Data" in den Neurowissenschaften ein Wandel von vielen individuellen, "vertikalen" Bemühungen zur Problembehandlung von einzelnen Fragestellungen hin zu generalisierten, "horizontalen" Bemühungen, durch die Akkumulation und Aggregation von gesammelten Daten, geführt [29]. Dieser Trend ist sowohl im präklinischen als auch im klinischen Kontext erkennbar und zeigt, dass in Wissenschaftskreisen das Bestreben große Datenmengen transparent darzustellen vorhanden ist. Zur Unterstützung eines adäquaten Studiendesigns, Datenaustausches und einer retrospektiven Bewertung von klinischen Studien entwickelte das amerikanische *National Institute of Neurological Disorders and Stroke (NINDS)*, eine Abteilung des *National Institute of Health*,

sogenannte *Common Data Elements* (CDE's). Mithilfe der Definition von gängigen Datenelementen sollen relevante Studiendaten für Wissenschaftler einfacher auffindbar sein und eine rege Teilnahme an der Konzeption von klinischen Studien anregen [3]. Das Studiendesign und dessen Ausführung stehen im Mittelpunkt dieses Vorhabens. Allerdings wird das bestehende Hauptproblem der Informationsflut bei ausbleibender Identifikation erfolgversprechender Therapieansätze auf dem Niveau von präklinischen Studien nicht behandelt, so dass eine Fortführung in der klinischen Umgebung erst gar nicht zu Stande kommt. Klinische Konzepte, wie die Einhaltung von Qualitätsstandards, institutsübergreifende Kooperationen und der Datenaustausch gewinnen in der präklinischen Forschung stetig an Bedeutung [26]. Ein Ansatz, um präklinische und klinische Studien über TRV zu koordinieren, ist die Verwendung von Rohdaten einiger ausgewählter Labors in der VISION-SCI Datenbank, unter deren Anwendung direkte Vergleiche mit den durch die NINDS postulierten *Common Data Elements* die Translation von präklinischen in klinische Studien unterstützt wird [11]. Diese Herangehensweise hat den Vorteil, dass die Qualität der Daten sehr hoch und zuverlässig ist. Andererseits sind die verwendeten Datenmengen, im Vergleich zur vorherrschenden Informationsflut, sehr gering. Keins der oben beschriebenen Verfahren bietet praktikable und leicht anwendbare Lösungsansätze für einen angemessenen Informationsaustausch zwischen verschiedenen Interessengruppen. Eine einzige Anwendung, die alle an der präklinischen Forschung über TRV teilnehmenden Fachbereiche vereint, ist nicht vorhanden. Jede Interessengruppe entwickelt individuelle strategische Werkzeuge für einen vermeintlich effizienten Informationsaustausch, der allerdings nicht für alle involvierten Fachbereiche adäquat ist. Daher empfiehlt sich eine Plattform, die von allen Interessengruppen als Ausgangspunkt zur Förderung der individuellen Vorhaben in gleichen Anteilen verwendet wird. Eine zentrale Ontologie zur Repräsentation des Kernwissens aus der Domäne der TRV kann dieser Aufgabe gerecht werden. Darüber hinaus kann sie eine transparente, unmissverständliche Berichterstattung unterstützen.

1.4 Ziele der Arbeit

Wissenschaftliche Publikationen dienen als Kommunikation einer Arbeitsgruppe mit dem restlichen wissenschaftlichen Umfeld eines Fachbereiches. Es ist hier, dass sowohl das Studiendesign als auch die gewonnenen Erkenntnisse präsentiert und der

Wissensgemeinschaft mitgeteilt werden. Aufgrund der großen Anzahl an wissenschaftlichen Publikationen im präklinischen Forschungsgebiet über TRV empfiehlt sich die Identifikation und Formalisierung essenzieller Konzepte, um einen transparenten Wissensaustausch und die Reproduktion von präklinischen Studien zu ermöglichen. Zur Umsetzung dieses Vorhabens soll die SCIO (*Spinal Cord Injury Ontology*) für eine systematische Erfassung des Wissens über therapeutische Konzepte, relevante Parameter und Ergebnisse aus der reichlich vorhandenen präklinischen Literatur zum Thema TRV entwickelt werden. Die Bereitstellung einer Konzeptualisierung des Domänenwissens soll durch die Verknüpfung von relevanten Konzepten und Relationen mit Begriffen aus der natürlichen Sprache, wie sie in wissenschaftlichen Publikationen verwendet wird, erfolgen. Die Ontologie soll in ihrer Struktur dem logischen Aufbau einer wissenschaftlichen Publikation ähneln und relevante Daten konzis, zusammenhängend und übersichtlich darstellen. Dieses Vorhaben liegt nahe, da sich die meisten Studien zur Aufrechterhaltung der guten wissenschaftlichen Praxis in ihrem strukturellen Aufbau sehr ähneln. Daraus ergibt sich die Notwendigkeit Details über methodische Abläufe und Ergebnisse aus dem Volltext von Publikationen zu extrahieren. Konsekutiv sollen sie als Konzepte definiert und in das Gesamtkonstrukt der Ontologie integriert werden. Zur Gewährleistung einer angemessenen Anwendung der Ontologie müssen Richtlinien für die Sicherstellung einer adäquaten und uniformen Anwendung bei der Annotation von wissenschaftlichen Texten durch verschiedene Personen erstellt werden. Die annotierten Ergebnisse müssen durch Querverweise zu ihren Komponenten einen klaren Einblick in den experimentellen Aufbau und das Durchführungsprotokoll ermöglichen, damit eine Evaluation präklinischer Studien erleichtert wird. Zur Validierung ihrer Effizienz sollte die Ontologie bei ihrer Anwendung zur manuellen Annotation und Datenextraktion von Texten anwendbar sein und ein hohes Maß an Übereinstimmung zwischen verschiedenen Benutzern aufweisen. Die Applikation der Ontologie auf wissenschaftliche Texte zum Thema TRV soll die Extraktion und Offenlegung signifikanter Ergebnisse aus präklinischen Studien vereinfachen, damit deren Umsetzung in eine klinische Studie zeitgerecht evaluiert und auf ihre Sinnhaftigkeit überprüft werden kann. Die Ontologie sollte durch ihre Anwendung Diskrepanzen bei der Berichterstattung in wissenschaftlichen Publikationen offenlegen. So können aus der Zusammenfassung des experimentellen Aufbaus und der annotierten Ergebnisse Rückschlüsse auf die Aussagekraft einer Studie gezogen werden. Die Ontologie soll

eine objektive Bewertung von einzelnen präklinischen Studien und den objektiven Vergleich zwischen verschiedenen präklinischen Studien erlauben. Prospektive Ziele, die im weiteren Verlauf über diese Arbeit hinaus gefördert werden können, sind (1) die Erweiterbarkeit der Ontologie als Grundlage und Rückgrat für die automatische Datenextraktion, indem ihre Konzepte und Relationen in eine Datenextraktionssoftware integriert werden, um in einer externen Validierung die automatische Datenextraktion von Publikationen und (2) die Zentralisation von aus präklinischen Studien gewonnenen Daten in eine für Wissenschaftler frei zugängliche Datenbank zur Analyse der gesamten vorhandenen Evidenz aus präklinischen Studien über TRV zu ermöglichen. Das langfristige Ziel der SCIO ist die Vereinfachung und Beschleunigung einer zeitgerechten Translation von erfolgversprechenden präklinischen Studien in klinische Studien.

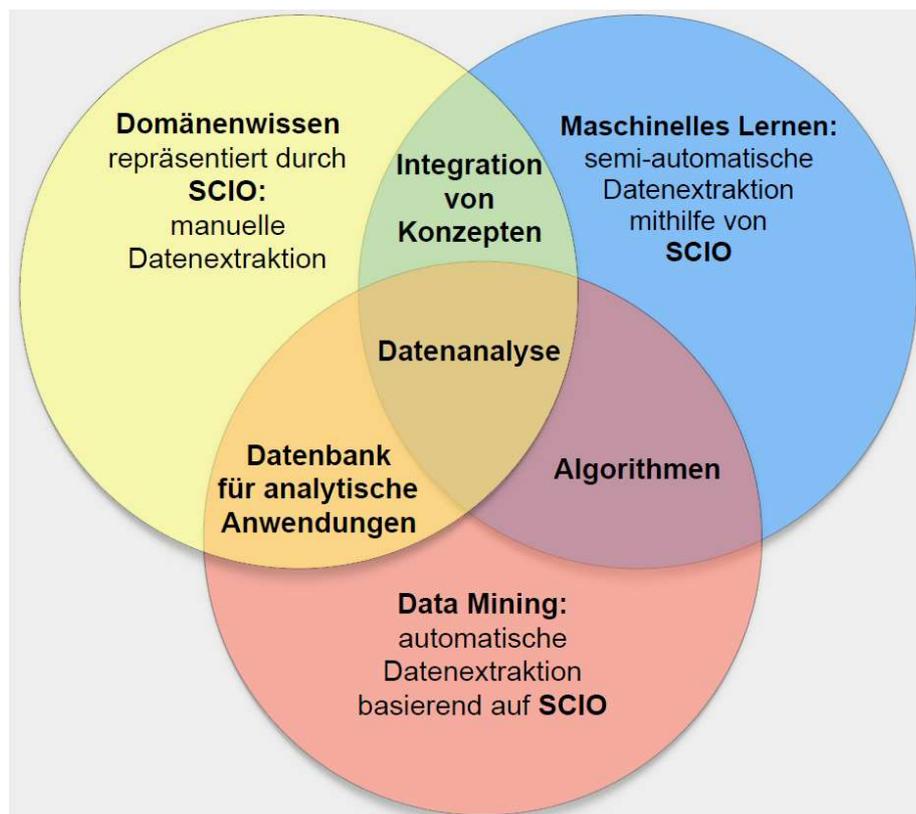


Abb. 1: **Anwendungsbereiche und Erweiterbarkeit der SCIO.** Die Ontologie steht im Zentrum der Weiterentwicklung von Strategien zur Förderung eines offenen Wissensaustausches.

Die komplexen Beziehungen zwischen der Ontologie und ihrer erweiterbaren Anwendungsbereiche gehen aus Abb. 1 hervor. Die Erstellung der Ontologie wird durch die gelbe Sphäre repräsentiert. Die blaue Sphäre zeigt die Weiterentwicklung und

Validierung der Ontologie durch die Anwendung ihrer Konzepte bei der semi-automatischen Datenextraktion bis hin zum *Data Mining* (rote Sphäre) für eine vollautomatische Datenextraktion. Die sich überlappenden Bereiche der Sphären heben die Interaktionen zwischen den unterschiedlichen Kompetenzbereichen hervor. Bei den Interaktionen handelt es sich um iterative Prozesse, die für die Aktualisierung der Ontologie notwendig sind. Hieraus wird ersichtlich, dass der Einfluss der Ontologie für die automatisierte Akquirierung von Datensätzen eine Voraussetzung ist. Die SCIO soll die Rückgratstruktur für die unterschiedlichen Formen der Datenextraktion und der Konzeptualisierung des Domänenwissens über traumatische Rückenmarksverletzungen in präklinischen Tierstudien bilden. Sie soll der Ausgangspunkt und das Zentrum für eine adäquate Repräsentation des Domänenwissens und dessen Verarbeitung zur Evaluation von Therapieansätzen und des experimentellen Aufbaus darstellen. Die Reichweite der SCIO geht aus ihren vielen Einsatzbereichen hervor und suggeriert eine Verbesserung beim Zugriff auf ein evidenzbasiertes Wissen, was für die Reproduktion von präklinischen Studien und retrospektiven Analysen zur Therapieentwicklung als Unterstützung dienen soll.

2. Material und Methoden

Für die Festlegung des Aufgabenbereiches der Ontologie mussten ihre Eigenschaften zunächst vergegenwärtigt werden. Im Allgemeinen repräsentiert eine Ontologie die charakteristische Terminologie, die wesentlichen Konzepte (mit ihren Klassifikationen und Relationen) und die Axiome einer Domäne [12]. Angewandt bedeutet dies, dass zur Diskussion von Themen in einer Domäne D unter Anwendung einer Sprache S , die Ontologie einen Katalog von verschiedenen Gattungen bereitstellt, deren Existenz in der Domäne D angenommen wird und durch Konzepte mitsamt ihren Relationen in der Sprache S repräsentiert werden können [12]. Die Ontologieentwicklung erfolgte in Anlehnung an die empfohlene Vorgehensweise der Methontology [21]: (1) Präzisierung des Leistungsumfanges und des Anwendungsbereiches der Ontologie basierend auf dem aktuellen Kenntnisstand in der Domäne der TRV, (2) Akquisition von Wissen aus wissenschaftlichen Texten, (3) Konzeptualisierung der in natürlicher Sprache verwendeten Terminologie, (4) eventuelle Integration von wiederverwendbaren Definitionen aus anderen Ontologien, (5) Implementierung der Ontologie, (6) Evaluation und (7) Dokumentation. Bei den vorgeschlagenen Schritten für das *Ontology*

Engineering ist eine strenge Einhaltung der Reihenfolge nicht notwendig [21], da es sich hierbei um einen individuellen Prozess handelt. Somit können die oben definierten Schritte parallel oder zeitlich versetzt erfolgen. Bevor der Entwicklungsprozess der SCIO begann, mussten zunächst ihre Ziele und ihr Kompetenzbereich festgelegt werden.

2.1 Definition des Kompetenzbereiches

Zur Bestimmung des Inhalts sowie des Umfangs der Wissensrepräsentation durch die Ontologie erfolgte eine Analyse der Domäne und ihrer Hauptmerkmale, um das Paradigma in der präklinischen Forschung über TRV und die Vorgehensweise bei der experimentellen Ausführung zu erfassen. In diesem Zusammenhang lag das Hauptaugenmerk auf der Beurteilung von pharmazeutischen Wirkstoffen und deren Kapazität eine Genesung nach induzierter traumatischer Rückenmarksverletzung in präklinischen Studien herbeizuführen [20]. Das Ziel der Ontologie ist die systematische Erfassung von therapeutischen Konzepten, die Beschreibung relevanter Parameter und die Präsentation der Ergebnisse aus präklinischen Tierstudien über traumatische Rückenmarksverletzungen. Da diese Konzepte in wissenschaftlichen Publikationen aufzufinden sind, diente deren Aufbau als Vorlage für die Entwicklung der Ontologie. Der eigentliche Aufbau von wissenschaftlichen Publikationen setzt sich aus zwei Hauptkomponenten zusammen. Zum einen ist der Aufbau durch das typische Format der schriftlichen Berichterstattung, wie es in wissenschaftlichen Veröffentlichungen aufzufinden ist und zum anderen durch den eigentlichen experimentellen Aufbau gekennzeichnet. Die schriftliche Berichterstattung erfolgt gemäß den Statuten der guten wissenschaftlichen Praxis durch die Beschreibung des wissenschaftlichen Hintergrundes, ggf. als Teil einer Einleitung, durch den Material- und Methodenteil, durch die Darlegung der Ergebnisse und durch einen analytischen Diskussionsabschnitt. Das typische präklinische *study design* wird durch den experimentellen Aufbau beschrieben, der sich aus den verschiedenen experimentellen Tiergruppen, durch eine oder ggf. mehrere experimentell induzierte Rückenmarkverletzungen, gefolgt von Untersuchungsmethoden, z.B. histologische Untersuchungen oder funktionale Bewegungsanalysen, und durch die Präsentation der Ergebnisse zusammensetzt. Beide Hauptkomponenten beinhalten relevante Konzepte, die für eine adäquate Repräsentation der Domäne und für die Beurteilung experimenteller Ergebnisse notwendig sind. In diesem Zusammenhang spielen alle Komponenten einer präklinischen Studie sowohl im

Einzelnen als auch in ihrer Kombination eine essenzielle Rolle. Zur Festlegung einer grundlegenden Struktur wurde der allgemeine experimentelle Aufbau und Ablauf einer präklinischen Studie als Vorlage für die Ontologie genommen. Diese Strategie sollte das typische wissenschaftliche Arbeiten und Denkschema in der präklinischen Forschung auf die Ontologie übertragen.

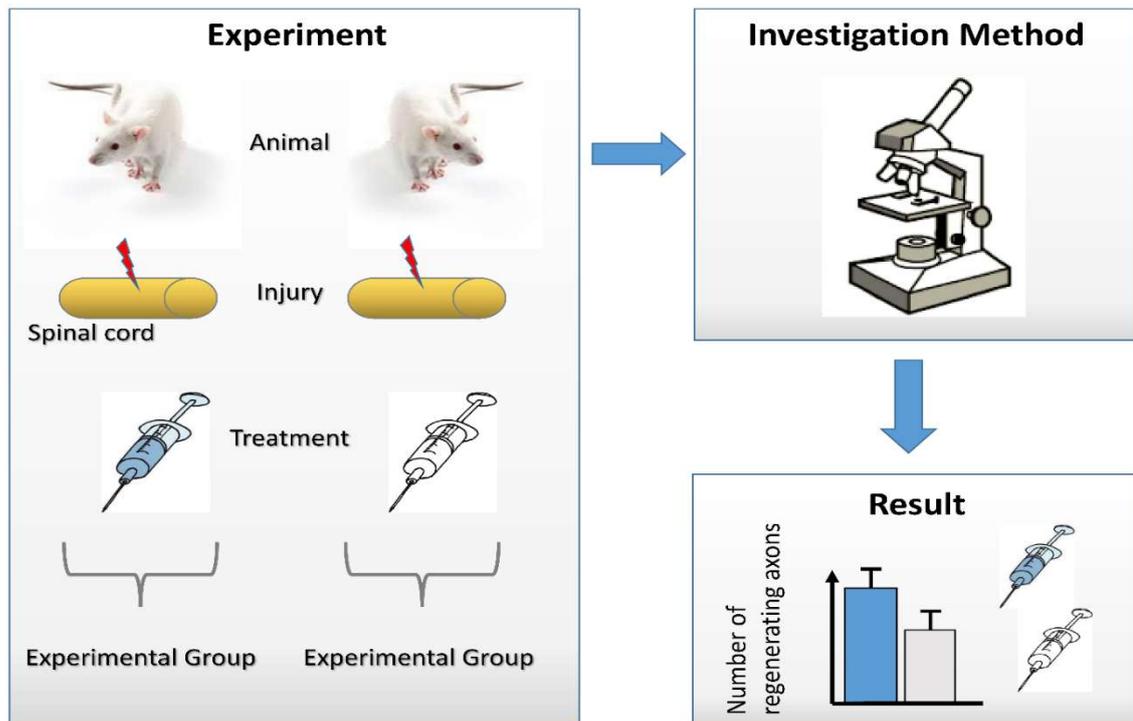


Abb 2. **Schematischer Überblick der grundlegenden Logik einer präklinischen Studie.** [5] Die Effizienz der Datenextraktion mit der Ontologie soll durch ihre Anlehnung an den strukturellen Aufbau einer wissenschaftlichen Publikation gewährleistet werden.

Abb. 2 zeigt die drei Hauptabschnitte einer wissenschaftlichen Publikation über TRV bestehend aus sechs Hauptkonzepten, die für das Verständnis der logischen Konfiguration einer präklinischen Studie notwendig sind. Beim typischen experimentellen Ablauf wird zunächst bei einer festgelegten Anzahl von Tiergruppen eine Rückenmarksverletzung experimentell induziert. Wahlweise erfolgt bei den experimentellen Gruppen eine konsekutive medikamentöse Therapie oder keine Therapie. Die therapeutischen Anwendungen auf das verletzte Rückenmark werden durch investigative Methoden evaluiert. Abschließend werden die Ergebnisse der unterschiedlichen experimentellen Gruppen miteinander verglichen und analysiert. Das analytische Lesen von mehr als fünfzig wissenschaftlichen Publikationen zeigte, dass sich ihr Aufbau als Rückgratstruktur für die Ontologie prädestinierte, um eine adäquate

Datenextraktion zu gewährleisten. Anhand des oben beschriebenen experimentellen Ablaufs einer präklinischen Studie wurden bereits vor der Ontologieentwicklung informelle Kompetenzfragen [14] zur Unterstützung bei der Festlegung ihrer Grundstruktur und ihres aus wissenschaftlichen Publikationen einzufangenden Informationsgehaltes gestellt. Folgende Fragen mussten durch die Ontologie beantwortet werden:

1. Wie ist eine präklinische Studie aufgebaut?
2. Wie setzten sich die experimentellen Gruppen zusammen?
3. Wie und zu welchem Zeitpunkt werden durch bestimmte Verletzungsmodelle Rückenmarkverletzungen induziert?
4. Bei welchen experimentellen Gruppen wurden welche Therapien zu welchem Zeitpunkt durchgeführt?
5. Welche Untersuchungsmethoden werden zur Beurteilung der Therapien angewendet?
6. Zeigen sich beim Vergleich der unterschiedlichen experimentellen Gruppen signifikante Veränderungen?

Über den experimentellen Aufbau und Ablauf hinaus verliehen die informellen Kompetenzfragen einen initialen Eindruck über das Format der Berichterstattung und der Darstellung der oben beschriebenen Hauptkonzepte, die in wissenschaftlichen Publikationen regelmäßig verwendet werden. Sie bildeten den Rahmen der Gesamtkonstruktion der Ontologie, der weitere spezifische Konzepte aus der Domäne der TRV umfassen sollte. Nachdem die Gerüststruktur der SCIO zunächst definiert wurde, mussten formelle Kompetenzfragen zur genaueren Zielsetzung der Ontologie gestellt und beantwortet werden. Diese Fragen sollten die notwendige Evidenz und Beurteilbarkeit von präklinischen Studien und die Nachhaltigkeit der Ontologie fördern. Die formellen Kompetenzfragen zielen bereits auf relevante Aspekte, deren Aussagekraft Aufschluss darüber geben soll, ob die vorliegenden Studien erfolgsversprechende Therapieansätze beinhalten und ob ein ausreichendes Maß an Information vorhanden ist, um ein beschriebenes Experiment replizieren und eventuell in der klinischen Umgebung weiterführen zu können. Demnach soll die Ontologie folgende Kompetenzfragen beantworten können [3]:

1. Welche Therapien haben zu positiven Ergebnissen bei unterschiedlichen Tierspezies geführt?
2. Welche Therapien haben zu positiven Ergebnissen sowohl bei funktionalen als auch bei nicht funktionalen Tests geführt?
3. Welche Therapien zeigen nur positive Resultate bei Teil-Läsionen, aber nicht bei Voll-Läsionen?
4. Welche Therapien zeigen einen funktionalen Effekt sowohl bei thorakalen als auch bei zervikalen Läsionen des Rückenmarks?
5. Nach wie vielen Wochen zeigt sich bei starken Kontusionen des thorakalen Rückenmarks eine Verbesserung der funktionalen Bewegungseinschränkung?
6. Welche investigativen Verfahren zeigen zum frühesten Zeitpunkt Unterschiede zwischen behandelten Tiergruppen und Kontrollgruppen bei funktionalen und nicht funktionalen Tests?

Aus den formellen Kompetenzfragen geht hervor, dass die Ergebnisse und deren Rückschlüsse auf die therapeutischen Verfahren im Fokus der Repräsentation des Domänenwissens liegen. Beim Vergleich zwischen den formellen und informellen Fragen können Relationen zwischen Ergebnissen, Therapien, Verletzungsmodellen und experimentellen Gruppen erstellt werden und somit Auskünfte über den gesamten Aufbau und Ablauf einer präklinischen Studie bereitgestellt werden.

2.2 Konstruktion der Ontologie

Die Konstruktion der SCIO bedeutete, dass eine neue Ontologie erstellt wurde, die unabhängig von anderen Ontologien die Domäne der TRV auf dem Niveau des experimentellen Aufbaus und der gängigen Berichterstattungsformate repräsentieren soll. Hierbei handelt es sich um ein völlig neues Konzept, das seine Berechtigung in der Notwendigkeit für eine neue Herangehensweise bei der Kategorisierung und adäquaten Präsentation des vorhandenen Fachwissens hat. Darüber hinaus soll für Wissenschaftler eine angemessene Zugänglichkeit auf dieses Fachwissen ermöglicht werden. Nachdem die Kompetenzfragen postuliert wurden, erfolgte eine Analyse von 50 wissenschaftlichen Publikationen, um eine Übersicht und ein besseres Verständnis für die Schreibweise, den experimentellen Aufbau und die Schwerpunkte, respektive die Hauptkonzepte zu erlangen. Die Analyse von wissenschaftlichen Texten erfolgte unter

stetiger Berücksichtigung der informellen und formellen Kompetenzfragen, so dass bereits in der Planungsphase beim Lesen der Texte versucht wurde die gestellten Kompetenzfragen anzuwenden. Zunächst wurden durch handschriftliche Dokumentationen wichtige Terminologien und Satzabschnitte aus diesen Texten extrahiert, um die Semantik in ihrer Anwendung zu verstehen und um bereits relevante Konzepte aus der Domäne zu akquirieren. Diese Konzepte wurden gesammelt und in einer Excel Tabelle aufgelistet. Da die weltweit veröffentlichten Publikationen über TRV ausschließlich in englischer Sprache vorliegen, sollte die Domäne ebenfalls durch die *Spinal Cord Injury Ontology* in englischer Sprache repräsentiert werden. In Hinblick auf die Konstruktion einer Ontologie sollten die gesammelten Konzepte Teil der späteren Struktur der Ontologie werden, wohlwissend dass die Entwicklung der Ontologie ein dynamischer Prozess ist, der stets zu weiteren Veränderungen durch das Hinzufügen von weiteren Konzepten und Attributen führen konnte. Nach der tabellarischen Zusammenfassung der gesammelten Hauptkonzepte des präklinischen Forschungsbereiches wurde bereits vor Beginn der eigentlichen Entwicklungsphase festgelegt, wie die Ontologie im weiteren Verlauf validiert werden sollte. Die Determinierung potenzieller Validierungsmethoden erfolgte im Vorfeld, damit sie auf der Grundlage des strukturellen Aufbaus der Ontologie appliziert werden konnten. Zur Evaluation möglicher Validierungsmethoden entstand eine Kooperation zwischen dem CNR (*Center for Neuronal Regeneration*) einer externen Abteilung des Labors für Molekulare Neurobiologie der Neurologischen Klinik an der Heinrich-Heine-Universität Düsseldorf und dem CITEC (*Cluster of Excellence Cognitive Interaction Technology*) an der Universität Bielefeld. Hier wurden in drei getrennten Konferenzen im Rahmen einer fachübergreifenden Zusammenarbeit durch Power Point Präsentationen die Motivation und Ziele für die *Spinal Cord Injury Ontology* dargestellt und mögliche Methoden zur Validierung der Ontologie im weiteren Verlauf diskutiert. Hierbei wurde festgelegt, dass die Ontologie durch unterschiedliche Anwender ein hohes Maß an Übereinstimmung garantieren muss und dies durch die Anwendung von statistischen Ermittlungsverfahren dargestellt werden sollte. Außerdem sollte sie als Vorlage für das *Machine Learning* dienen, eine technische Anwendung, die nach der abgeschlossenen Entwicklung der Ontologie ihre Anwendung bei der automatischen Datenextraktion finden sollte. Nachdem die aktuellen und weiterführenden Kompetenzbereiche der Ontologie festgelegt wurden, begann die eigentliche Konstruktion der Ontologie.

2.2.1 Datenextraktion aus wissenschaftlicher Literatur

Durch die Analyse von wissenschaftlichen Publikationen wurden für die Entwicklung der Ontologie relevante Konzepte identifiziert. Basierend auf den von Kwon et al. [14] vorgeschlagenen Gesichtspunkten wurde ein Testkorpus von 122 Publikationen im *Portable Document Format* (PDF) aus einem Gesamtkorpus, bestehend aus 11.400 Publikationen der Wissenschaftsplattform PubMed, akkumuliert. Die für die Ontologieentwicklung festgelegten Auswahlkriterien schlugen vor, dass die ausgewählten Publikationen (1) ein *in vivo* Tiermodell beinhalten mussten, (2) dass das Rückenmark traumatisch verletzt wurde, (3) dass die Therapie über den systemischen Kreislauf verabreicht wurde und (4) dass es sich bei den Publikationen um *peer-reviewed publications* handelte [9]. Humanstudien wurden für die Entwicklung der SCIO exkludiert und nicht in Betracht gezogen. Die 122 ausgewählten Publikationen stimmten mit den Auswahlkriterien überein und stellten den initialen Testkorpus zur manuellen Extraktion von relevanten Konzepten dar. Die manuelle Extraktion und Kategorisierung von domänenspezifischen Konzepten erfolgten mit einer Excel-Datenbank. Sie dienten als Vorläufer für Klassen, die in das Konstrukt der Ontologie inkorporiert werden sollten. Die ursprüngliche Anzahl der Publikationen im Testkorpus erhöhte sich im weiteren Verlauf von 122 auf 140 Publikationen, um weitere Konzepte für eine allumfassende Repräsentation der Domäne durch die Ontologie sicherzustellen. Die Kompetenzfragen wurden in regelmäßigen Abständen auf die extrahierten Daten angewendet, um den notwendigen Wissensumfang für eine angemessene Repräsentation der Domäne durch die SCIO zu überprüfen. Das Resultat der Metaanalyse von 140 Publikationen war die Extraktion von über 800 Datenpunkten, die Aufschluss über das *experimental design*, über individuelle Ergebnisse von verschiedenen experimentellen Gruppen, und über die Tier-, Verletzungs- und Therapiemodelle gaben. Mit den manuell extrahierten Informationen ließen sich die informellen Kompetenzfragen bereits beantworten, allerdings lagen die extrahierten Entitäten zu diesem Zeitpunkt lediglich unabhängig und in keinem Zusammenhang mit der restlichen Information einer Publikation vor, so dass die Beantwortung der formellen Kompetenzfragen noch nicht möglich war.

2.2.2 Konzeptualisierung und Formalisierung der Domäne

Die aus wissenschaftlichen Publikationen extrahierten Datenpunkte, mussten zunächst in die Struktur einer Ontologie integriert werden. Eine Ontologie entspricht einer Sprache bestehend aus Definitionen und Einschränkungen von Axiomen durch die formale Beschreibung von Objekten mitsamt ihren Eigenschaften und Relationen [14]. Das Grundgerüst der Ontologie sollte einer Einteilung aus Oberklassen und den dazugehörigen Ausprägungen von Unterklassen (*subclasses*, *instances*), auch Objekte genannt, entsprechen. Jeder Klasse sollten Eigenschaften (*attributes*, *slots*) zugeordnet werden, um ein Relationsgefüge innerhalb der individuellen Klassen herzustellen. Bevor die eigentliche Konstruktion der Ontologie begann, musste für ein allgemeines Verständnis der Grundlogik einer präklinischen Studie ein Überblick der extrahierten Hauptkonzepte und ihrer Relationen zueinander verschafft werden. Zur bildlichen Darstellung der Hauptkonzepte der Ontologie, einschließlich ihrer Beziehungen zueinander, wurde ein Klassendiagramm im *Unified Modeling Language* (UML) erstellt.

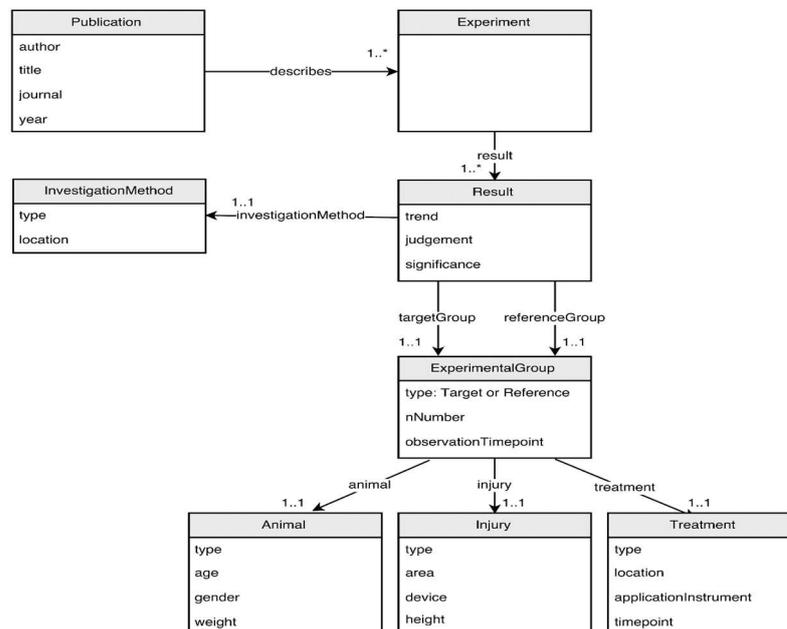


Abb. 3 Vereinfachte Darstellung des logischen Aufbaus eines präklinischen Experimentes in der Domäne der traumatischen Rückenmarksverletzungen im Unified Modeling Language (UML). Die in Grau unterlegten Hauptklassen bestehen aus der Summe der ihr untergeordneten Unterklassen. Ein Experiment beschreibt ein oder mehrere Resultate, die Aufschluss über die angewendeten Untersuchungsmethoden auf eine experimentelle Gruppe geben. In diesem Zusammenhang beschreibt die Hauptklasse Experimental Group alle Entitäten des experimentellen Aufbaus und Ablaufs.

Bei der manuellen Datenextraktion aus dem Testkorpus konnte nach sorgfältiger Analyse der wissenschaftlichen Literatur anhand des experimentellen Aufbaus und Ablaufs festgestellt werden, dass die Ontologie aus acht Hauptklassen bestehen musste, um die Domäne auf dem Niveau einer wissenschaftlichen Publikation in ihrer Gesamtheit repräsentieren zu können. In Abb. 3 werden die acht Hauptklassen der SCIO Ontologie im *Unified Modeling Language* (UML) dargestellt: (1) *Publication*, (2) *Experiment*, (3) *Experimental Group*, (4) *Animal*, (5) *Injury*, (6) *Treatment*, (7) *Investigation Method* und (8) *Result*. Bei den in Abb. 3 in Grau unterlegten Überschriften handelt es sich um die acht Hauptklassen. Die in den unterschiedlichen Rechtecken unterhalb der Klassen aufgeführten Bezeichnungen sind die individuellen Attribute einer Klasse (*attributes/ slots*). Eine Publikation kann ein oder mehrere Experimente beinhalten, die zu einem Ergebnis oder mehreren Ergebnissen führen können. Ein Ergebnis beschreibt das Resultat eines Tests basierend auf der Observation eines Vergleichs zwischen einer experimentellen Zielgruppe (*Target Group*) mit einer experimentellen Referenzgruppe (*Reference Group*). Ein Ergebnis bezieht sich auf genau ein investigatives Verfahren, das auf den Vergleich von zwei experimentellen Gruppen appliziert wurde. Eine experimentelle Gruppe (*Experimental Group*) besteht aus einer bestimmten Anzahl von Tieren, die einer identischen Verletzung ausgesetzt wurden und die gleiche Therapie bekamen. Die experimentelle Gruppe ist ein Aggregat aus den drei Hauptklassen *Animal*, *Injury* und *Therapy* und gibt Auskunft über die Art der Therapie, die nach experimentell verursachter Verletzung bei einer bestimmten Tierspezies angewendet wurde. Diese Anordnung der Klassen erlaubt es willkürliche Vergleiche zwischen unterschiedlichen experimentellen Gruppen bei einer vorgegebenen Untersuchungsmethode (*Investigation Method*) im Rahmen einer bestimmten Studie durchzuführen. So kann eine Kontrollgruppe mit einer Anzahl N von Ratten mit der experimentellen Gruppe 1 verglichen werden, die nach einer Kontusion der Lendenwirbel L2 bis L4 das Präparat A intravenös verabreicht bekam, oder sie könnte mit der experimentellen Gruppe 2 verglichen werden, die im selben Abschnitt des Rückenmarks nach vorausgegangener Kontusion das Präparat B intravenös verabreicht bekam. Zusätzlich können die experimentellen Gruppen 1 und 2 miteinander verglichen werden (z.B. behandelte *Target Group 1* versus *Control Group*, behandelte *Target Group 2* versus *Control Group* und *Target Group 1* versus *Target Group 2*). Hierbei handelt es sich jeweils um paarweise Vergleiche, die unabhängig voneinander, je einem Ergebnis und somit insgesamt drei Ergebnissen entsprechen. In

diesem Fall sollen die drei Ergebnisse aus einem Experiment unter Wahrung ihrer individuellen Eigenschaften durch die SCIO identifiziert werden. Darüber hinaus ist nicht nur ein Vergleich von zwei experimentellen Gruppen miteinander möglich. Es können auch mehrere Gruppen in unterschiedlichen Konstellationen miteinander verglichen werden. Dies könnte zutreffen, wenn der Autor eine statistische Clusteranalyse von mehreren experimentellen Gruppen, die bei gleicher Therapie zu verschiedenen Zeitpunkten nach vorausgegangener Verletzung des Rückenmarks mit einer Kontrollgruppe verglichen werden sollen (z.B. *Target Group 1*, *Target Group 2*, *Target Group 3* versus *Control Group*). Dieser Vergleich entspricht einem Ergebnis. Im Paradigma der SCIO-Ontologie ist die Anzahl der Ergebnisse, die aus einer präklinischen Studie hervorgehen proportional zu den einzelnen Vergleichen zwischen experimentellen Gruppen, respektive zwischen einer *Target Group* mit einer *Reference Group* oder zwischen mehreren *Target Groups* mit einer *Reference Group* und umgekehrt, zwischen einer *Target Group* mit mehreren *Reference Groups*. Zusammenfassend kann nach dem analytischen Lesen und der Extraktion von domänenspezifischen Konzepten konstatiert werden, dass das Ergebnis (*Outcome*) einer Studie im Mittelpunkt einer präklinischen Publikation über TRV steht. Das Aggregat der experimentellen Gruppen und die investigativen Methoden stehen unabhängig voneinander in direktem Zusammenhang mit dem Ergebnis. Jedes Ergebnis entspricht einem Vergleich zwischen unterschiedlichen experimentellen Gruppen und beinhaltet Informationen über die Zusammensetzung der Tiergruppen, der induzierten Rückenmarkverletzung und der Therapie. Zur Planung einer zusammenhängenden Repräsentation und Anwendung dieser Konzepte mussten die in Abb. 3 dargestellten Klassen mit all ihren Objekten und Attributen kategorisiert werden. Eine Technik zur einfachen Repräsentation von zusammenhängenden Wissenseinheiten sind sogenannte *Object-Attribute-Value Triplets* (OAVTs), die zur Darstellung individueller Klassen mitsamt ihren Attributen und Werten von semantischen Netzen verwendet werden [12]. Beim analytischen Lesen von wissenschaftlichen Publikationen wurden wichtige Merkmale zwischen verschiedenen Objekten identifiziert und in einen logischen Zusammenhang gestellt. Hierfür eignete sich die Applikation von OAVTs für die Vergegenwärtigung und Vernetzung aller relevanten Hauptkonzepte mitsamt ihren Eigenschaften. Die OAVTs unterstützen die Erstellung von kurzen und zugleich signifikanten Beziehungen zwischen Klassen und ihren Eigenschaften. Ihre Anwendung ermöglichte die genaue Definition von zahlreichen Charaktereigenschaften, die aus

einer Klasse hervorgehen können und für das Verständnis des experimentellen Aufbaus und Ablaufs von großer Bedeutung sind.

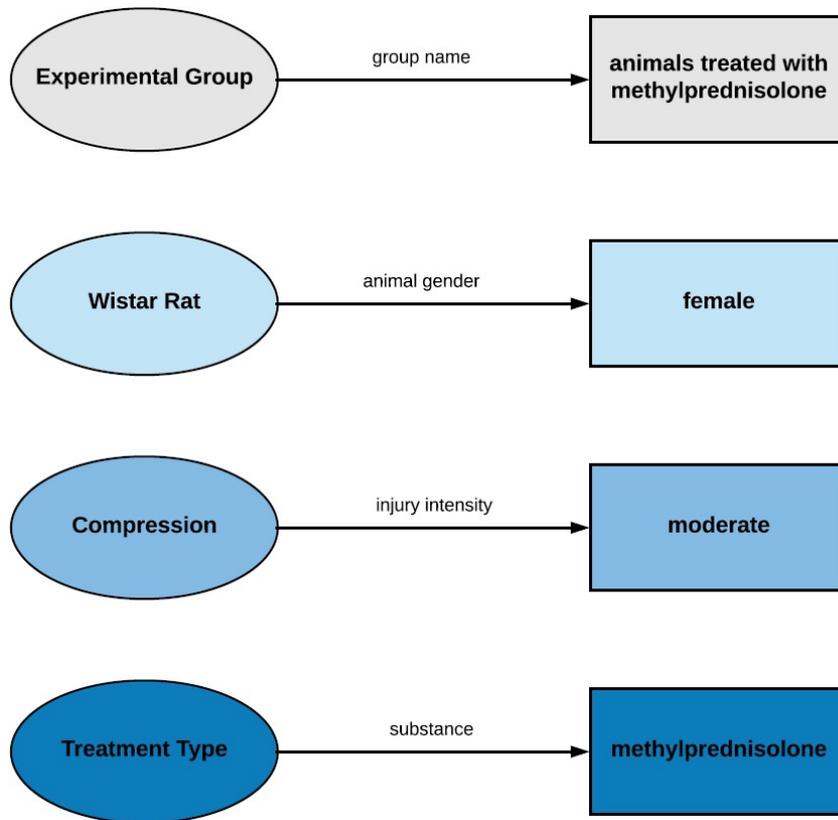


Abb. 4 Vier Object-Attribute-Value Triplets einer experimentellen Gruppe.

Abb. 4 zeigt vier untereinanderstehende Beispiele für ein *Object-Attribute-Value Triplet* (OAVT). Sie repräsentieren eine vereinfachte Form des allgemeinen experimentellen Ablaufs aus der Sicht einer experimentellen Gruppe durch die Darstellung von vier Klassen mit Ihren Eigenschaften und den dazugehörigen Werten. Über den von einer Klasse ausgehenden Pfeil steht das jeweilige Attribut. Der Pfeil zeigt auf den zum Attribut gehörenden Wert. Zusammenfassend zeigen sich in Abb. 4, von oben nach unten betrachtet, vier unabhängige OAVTs einer experimentellen Gruppe, die im englischen Fließtext einer Publikation als „Tiere, die mit Methylprednisolon behandelt wurden“ bezeichnet wird. Aus der Abbildung geht hervor, dass die experimentelle Gruppe aus weiblichen Wistar-Ratten besteht, deren Rückenmark einer moderaten Kompressionsverletzung ausgesetzt wurde, um konsekutiv durch die intravenöse Verabreichung von Methylprednisolon behandelt zu werden. Hierbei handelt es sich um ein vereinfachtes Beispiel, denn von einer Klasse können eine Vielzahl von Pfeilen (Attribute mit Werten) ausgehen. So könnte von der Klasse *Wistar-Rat* ein weiterer Pfeil mit dem Attribut *animal age* auf den Wert *adult* zeigen. Weitere Attribute könnten

die Anzahl der Tiere in einer experimentellen Gruppe, der genaue Abschnitt des verletzten Rückenmarks oder der genaue Zeitpunkt der Therapieanwendung sein. Diese Form der Darstellung, geprägt von einfachen, direkten Beziehungen zwischen Entitäten, erlaubten es zunächst alle wichtigen Hauptklassen und Unterklassen mit ihren Attributen zu definieren. Darüber hinaus konnten die beschriebenen Eigenschaften den Zusammenhang zwischen unterschiedlichen Klassen herstellen. Dieses Bottom-Up-Verfahren, respektive der auf höchster Detailebene beginnende Aufbau der Ontologie, hatte zur Folge, dass viele Klassen unabhängig zu anderen Klassen mit ihren Attributen in Beziehung gesetzt wurden, ohne dass ein zusammenhängender Überblick aus ihnen resultierte. Hierfür mussten zusätzliche Klassen und Attribute als Angelpunkte identifiziert werden, um die Verbindungen zwischen den unterschiedlichen Hauptklassen herzustellen und ein zusammenhängendes Wissenskonstrukt in Form der SCIO-Ontologie zu konzipieren. Einen dieser Angelpunkte stellte die experimentelle Gruppe dar. Anhand des logischen Aufbaus einer präklinischen Studie ging aus den Textanalysen hervor, dass im Material- und Methodenteil einer wissenschaftlichen Publikation die Komponenten einer experimentellen Gruppe durch drei Hauptklassen, die Tierspezies, die experimentell verursachten Verletzungen und die angewandte Therapie, definiert werden und im Zusammenhang eine große Einheit bilden. Diese Hauptklassen beschreiben alle für den experimentellen Aufbau und Ablauf relevanten Informationen und repräsentieren den gesamten Material- und Methodenteil einer wissenschaftlichen Publikation. Die experimentelle Gruppe stellt somit ein Aggregat aus drei essenziellen Hauptklassen mit den dazugehörigen Unterklassen und Attributen dar. Die Erstellung von semantischen Netzen diente der Darstellung und Konzeption von aus mehreren Hauptklassen bestehenden Aggregaten, um im Vorfeld die Zusammenhänge zwischen den unterschiedlichen Entitäten zu verstehen und um sie auf ihre Richtigkeit zu überprüfen. Mithilfe der semantischen Netze konnten die für sich alleinstehenden OAVTs miteinander in Zusammenhang gesetzt werden, um eine Übersicht der logischen Zusammenhänge zwischen den einzelnen Klassen und ihren Attributen herzustellen. Die semantischen Netze wurden zunächst bei der In-Bezugsetzung von Hauptklassen angewendet. Die den Hauptklassen untergeordneten Unterklassen wurden zur Gewährleistung einer verständlichen Transparenz außer Acht gelassen und würden erst im weiteren Verlauf bei der Entwicklung der Ontologie eine Rolle spielen. Die grafische Darstellung von semantischen Netzen ist nicht standardisiert, aber zur Entwicklung von Ontologien werden Klassen (*concepts*),

Objekte, als ein Teil von Klassen (*instances*), und die Eigenschaften (attributes/ slots) von Klassen mit ihren Werten verwendet [12]. Die Beziehungen zwischen Konzepten werden im *Unified Modeling Language* wie folgt beschrieben: "ist eine Art von" ("*a kind of*"), "ist Teil von" ("*is a part of*") und "ist ein/ eine" ("*is a*"). Zur Beschreibung von Beziehungen können auch andere, individuelle Bezeichnungen verwendet werden.

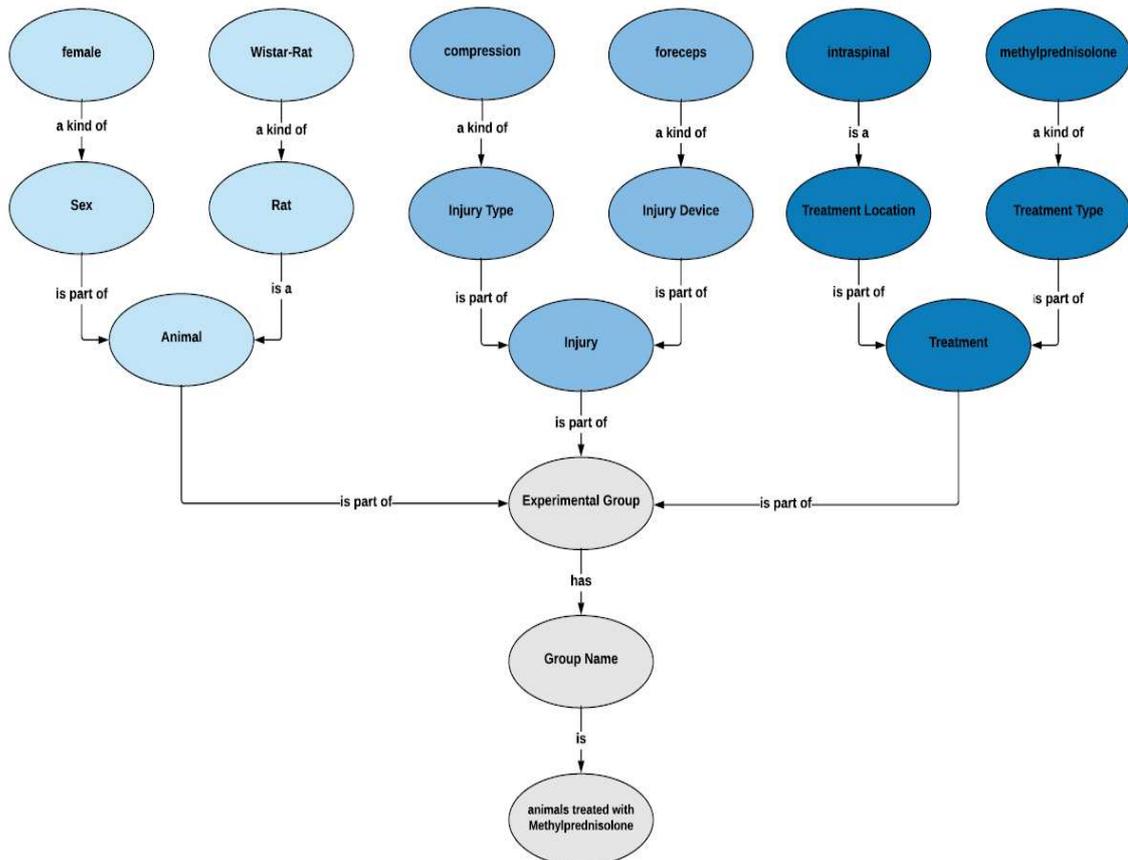


Abb. 5 **Das semantische Netz einer experimentellen Gruppe.** Es zeigt das Aggregat einer experimentellen Gruppe bestehend aus einer Tierspezies, der ihr zugefügten Verletzung und die applizierte Behandlungsmethode. Darüber hinaus wird die im Fließtext einer Publikation verwendete Bezeichnung der Gruppe angezeigt.

In Abb. 5 sind die zu einer experimentellen Gruppe gehörenden aggregierten Hauptklassen und deren Unterklassen zusammenhängend in einem semantischen Netz dargestellt und beschreiben alle Informationen, die für das Verständnis des experimentellen Ablaufs und allen dazugehörenden Entitäten von Bedeutung sind. Die Farbcodierungen stimmen mit denen aus Abb. 4 überein und sollen die Integration der OAVTs in das semantische Netz verdeutlichen. In Anlehnung an das psychologische Modell eines assoziativen Gedächtnisses sind semantische Netzwerke grafische Darstellungen von Konzepten, Objekten und Situationen einer Domäne, die durch

Relationen miteinander verbunden werden [12]. Aus der Grafik gehen die Hauptklassen *Animal*, *Injury* und *Treatment* nun als Unterklassen der *Experimental Group* hervor, die als Angelpunkt der ihr untergeordneten, aggregierten Klassen fungiert. Somit ist ein logischer Zusammenhang zwischen den Klassen hergestellt, der als Grundlage für den Entwurf der Ontologie dient. Weiterführend kann das semantische Netz den Ablauf bei der Annotation von wissenschaftlichen Texten mit Hilfe der Ontologie unterstützen. Neben den Hauptklassen sind einige Unterklassen aufgeführt, um das Ausmaß der mitgeführten Konzepte anzudeuten. Hier ist zu beachten, dass die Hauptklassen weitaus mehr Unterklassen beinhalten als in Abb. 5 dargestellt. Im Wesentlichen repräsentiert das in Abb. 5 gezeigte semantische Netz den Material- und Methodenteil, respektive den ersten Hauptabschnitt einer wissenschaftlichen Publikation. Die aus diesem Abschnitt einer wissenschaftlichen Veröffentlichung gewonnenen Daten werden im Ergebnis- und Diskussionsteil, dem zweiten Hauptabschnitt einer Publikation, bewertet und analysiert. Die *Experimental Group* vereint im Sinne eines Schlüsselkonzepts das Tier-, Verletzungs- und Therapiemodell einer präklinischen Studie. Darüber hinaus dient sie als Angelpunkt zwischen dem experimentellen Abschnitt und dem analytischen Abschnitt einer wissenschaftlichen Publikation. Zur Darstellung dieses Abschnittes mussten zusätzliche OAVTs definiert werden, um die essenziellen Zusammenhänge zwischen den experimentellen Gruppen und den aus dem experimentellen Ablauf entstandenen Ergebnissen herzustellen. Unabhängig vom Material- und Methodenteil werden im Ergebnis- und Diskussionsteil einer Veröffentlichung die verwendeten Untersuchungsmethoden zwischen bestimmten experimentellen Gruppen und die daraus entstehenden Beobachtungen und Resultate beschrieben. Die SCIO soll diese voneinander unabhängigen Abschnitte mit Hilfe von Schlüsselklassen verbinden, damit die zu einem Ergebnis gehörenden experimentellen Abläufe demselben zugeordnet werden können. In Abb. 6 ist die Hauptklasse *Investigation Method* in Grau unterlegt. Ähnlich wie die *Experimental Groups* des Material- und Methodenteils stellt die *Investigation Method* einen wichtigen Angelpunkt im Ergebnis- und Diskussionsteil dar. Basierend auf einer spezifischen Untersuchungsmethode, geht der Vergleich von zwei experimentellen Gruppen (in grau unterlegt) aus, die unterschiedlichen Therapien ausgesetzt wurden. Diese Vergleiche erfolgen im Ergebnisteil durch die Klasse *Outcome* und zeigen, wie der experimentelle Ablauf mit den Ergebnissen durch die Klasse *Investigation Method* verbunden wird. Eine Untersuchungsmethode ist die Voraussetzungen für einen Vergleich zwischen

unterschiedlichen experimentellen Gruppen. Sie bestimmt den thematischen Hintergrund auf dessen Basis unter anderem die motorischen Eigenschaften von Versuchstieren beurteilt werden sollen.

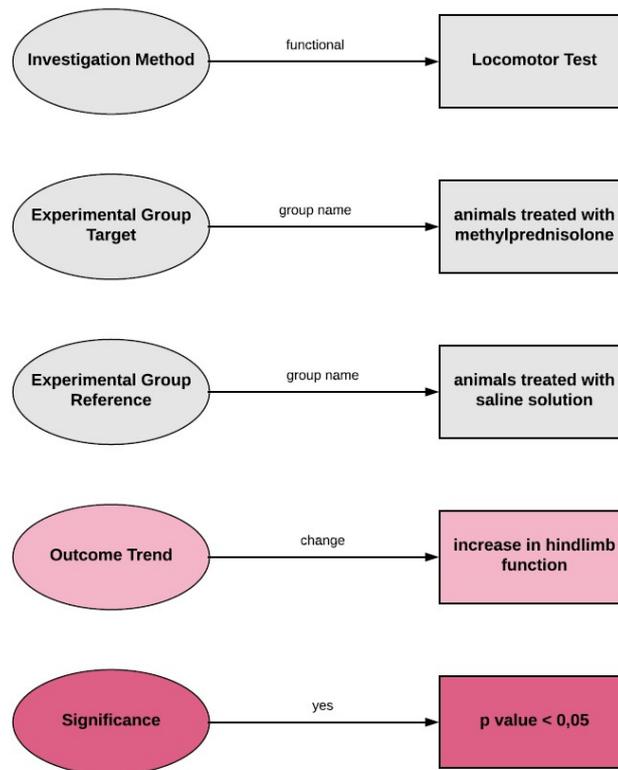


Abb. 6 **Fünf Object-Attribute-Value-Triplets eines Ergebnisses (Outcome)**. Wie zuvor sind in Abb. 6 auf der linken Seite in elliptischer Form Klassen zu erkennen, deren Attribute (Pfeile) auf einen Wert in rechteckiger Form zeigen. Die in grau gekennzeichneten Klassen *Investigation Method*, *Experimental Group Target* und *Experimental Group Reference* sind die drei Angelpunkte, die den Material- und Methodenteil mit dem Ergebnis- und Diskussionsteil verbinden.

Somit basiert das beschriebene Ergebnis auf die angewandte Untersuchungsmethode. Für jede Untersuchungsmethode und jeden damit verbundenen Vergleich zwischen experimentellen Gruppen entsteht ein *Outcome*. Abb. 6 zeigt fünf unabhängige OAVTs, die im Aggregat ein *Outcome* beschreiben. Unter Anwendung einer bestimmten Untersuchungsmethode zur Beurteilung der Lokomotion wurde die Funktion der Hinterbeine bei Tieren einer experimentellen Zielgruppe, mit der einer experimentellen Referenzgruppe verglichen. Hieraus ging hervor, dass bei der mit Methylprednisolon behandelten Zielgruppe im Vergleich zu einer mit Kochsalz behandelten Referenzgruppe eine Verbesserung bei der Beweglichkeit der hinteren Gliedmaßen festgestellt werden konnte. Zudem indiziert ein P-Wert von kleiner als 0,05, dass bei der

Zielgruppe eine signifikante Zunahme der Beweglichkeit an den Hinterbeinen festgestellt werden konnte.

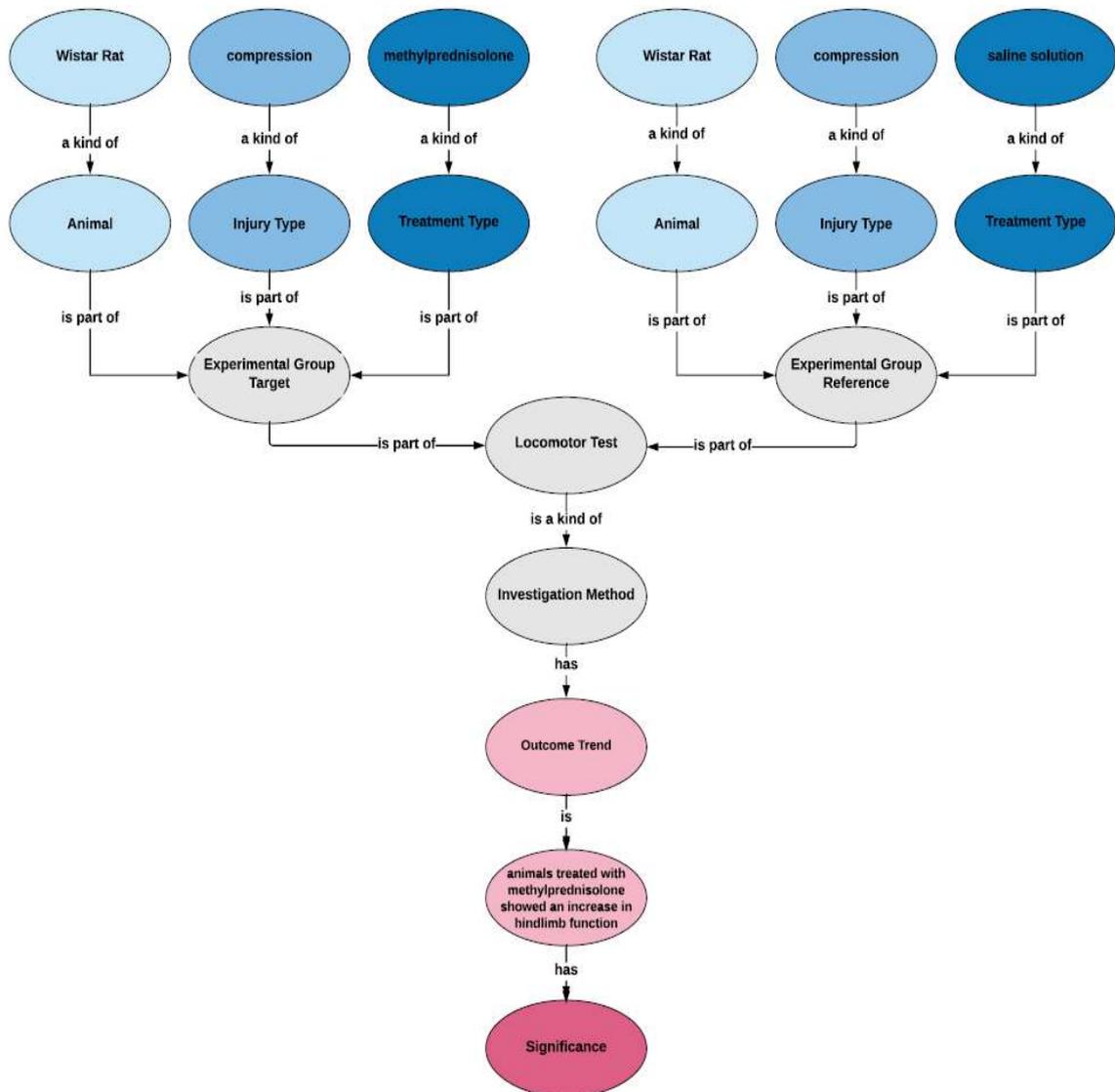


Abb. 7 **Das semantische Netz eines Ergebnisses.** Von oben nach unten werden alle Hauptkonzepte einer präklinischen Studie und der experimentelle Ablauf gezeigt. Aus dem experimentellen Aufbau (blau) und dem Vergleich zwischen einer experimentellen Zielgruppe und Referenzgruppe (grau) geht das Ergebnis (rot) hervor.

Im oberen Drittel von Abb. 7 sind in verschiedenen Blautönen die zu einer experimentellen Gruppe gehörenden Unterklassen erkennbar. Im mittleren Drittel zeigen sich die in Grau unterlegten Angelpunkte, die dem in blau gekennzeichneten experimentellen Ablauf, mit dem in unterschiedlichen Rottönen dargestellten Ergebnis verbinden. Die *Experimental Groups* des Material- und Methodenteils und die *Investigation Method* des Ergebnis- und Diskussionsteils setzen alle zu einem Ergebnis

gehörenden Klassen miteinander in Zusammenhang. Sie ermöglichen eine übersichtliche und zugleich transparente Präsentation des gesamten Wissens einer wissenschaftlichen Studie, indem alle Einzelheiten eines Ergebnisses beschrieben werden. Hierzu zählen erneut die Tiergruppen, die experimentell induzierten Verletzungen und die Behandlungsmethoden, die bei unterschiedlichen experimentellen Gruppen appliziert wurden und durch die Anwendung eines investigativen Verfahrens (*Investigation Method*) untersucht und anschließend durch ein Ergebnis (*Outcome*) beurteilt wurden. In den meisten Fällen geht eine Vielzahl von Ergebnissen aus einer präklinischen Studie hervor. Im Durchschnitt entstehen zehn voneinander unabhängige Ergebnisse, die die unterschiedlichen Behandlungsmethoden zu unterschiedlichen Zeitpunkten und Umweltbedingungen an verschiedenen Tiergruppen beschreiben. Das semantische Netz in Abb. 7 zeigt zur besseren Übersicht lediglich ein Ergebnis, gekennzeichnet durch einen Vergleich zwischen zwei experimentellen Gruppen, das neben einer Vielzahl von Ergebnissen bestehen kann. Im Hinblick auf die Beurteilung von präklinischen Studien spielt die Unterklasse *Significance* bei der Wissensmodellierung der Ontologie eine bedeutende Rolle. Sie soll bereits darüber Auskunft geben, ob eine präklinische Studie für eine erfolgreiche Weiterführung in einer klinischen Studie prädestiniert ist oder zumindest dafür in Erwägung gezogen werden kann. Bereits bei der Entwicklung der Ontologie wurde festgelegt, dass ein Ergebnis nur mit einem dazugehörigen P-Wert von kleiner als 0,05 als signifikant anerkannt und bei der späteren Annotation mithilfe der Ontologie annotiert werden sollte. Hierbei ist zu beachten, dass lediglich ein signifikantes Ergebnis *per se* nicht als Indikator für eine vermeintlich erfolgreiche Studie im Gesamten ausreicht. Hierzu wäre eine Analyse aller Ergebnisse mit den dazugehörenden Beurteilungen notwendig. Zudem ist das annotierte Ergebnis „*animals treated with methylprednisolone showed an increase in hindlimb function*“ der zum *Outcome Trend* gehörende Wert. Das Ergebnis wird durch eine Tendenz beschrieben, die entweder zu einer Zunahme (Unterklasse *Increase*) oder Abnahme (Unterklasse *Decrease*) eines Wertes führt. Das aus dem Fließtext annotierte Ergebnis zeigt, dass eine statistisch signifikante Zunahme (*Increase*) bei der Beweglichkeit der Tiere festzustellen war. In Abb. 7 werden nicht alle zu den Hauptklassen gehörenden Unterklassen aufgeführt, da es sich bei dem semantischen Netz um eine übersichtliche Darstellung von unterschiedlichen Relationsgefügen eines experimentellen Vergleiches handelt. Die Anwendung von semantischen Netzen ermöglichte die Beantwortung der informellen Kompetenzfragen

bereits vor der eigentlichen Konstruktion der Ontologie. Alle Aspekte der experimentellen Gruppen (*Animal*, *Injury* und *Treatment*) und der Untersuchungsmethoden gefolgt von der Signifikanz der Ergebnisse konnten festgestellt werden. Nachdem der typische Aufbau einer präklinischen Studie durch OAVTs und semantische Netze erörtert und definiert wurde, konnte mit der eigentlichen Konstruktion der Ontologie begonnen werden.

2.3 Der Aufbau der Spinal Cord Injury Ontology

2.3.1 Das Konzept der Klassenhierarchien

Für die Konzeption der Ontologie wurde die Protégé 3.3.1-Software verwendet, die plattformunabhängig auf den Betriebssystemen unterschiedlicher Computerhersteller anwendbar ist. Kennzeichnend für die Protégé-Software ist ihre benutzerfreundliche Plug-in-Architektur mit der eine Vielzahl von kleineren Programmen in das Hauptprogramm integriert und appliziert werden können. Eines dieser, für die Entwicklung der Ontologie, wichtigen Plug-ins ist Knowtator 3.3.1. Mit diesem Plug-in konnte durch die Anwendung der Ontologie die Annotation von wissenschaftlichen Texten erfolgen. Hierzu wurden wissenschaftliche Texte kopiert und formatgerecht im Knowtator-Plug-in abgespeichert und geöffnet, um dann im weiteren Verlauf annotiert zu werden. Vor der Annotation von Texten mussten die gewonnenen Erkenntnisse über den typischen Aufbau einer präklinischen Studie in die Ontologie integriert werden. Mithilfe der OAVTs und der semantischen Netze wurden zunächst die Hauptklassen und die dazugehörigen Unterklassen der experimentellen Gruppen mit Protégé modelliert. Die primäre Modellierung erfolgte durch die Zuordnung von Unterklassen zu komplementären Oberklassen. Für eine adäquate Repräsentation der Hauptklasse *Animal* wurden ihr Konzepte, wie die Tierspezies, das Alter, das Gewicht und das Geschlecht der Versuchstiere untergeordnet. Hieraus ergab sich eine Klassenhierarchie, die einer primären Ordnung von unterschiedlichen Hauptklassen und den dazugehörigen Unterklassen entsprach. Die sekundäre Ordnung der SCIO wurde durch die Anwendung von Attributen geschaffen, die eine Quervernetzung unterschiedlicher Hauptklassen ermöglicht. Die primäre Struktur der SCIO entspricht bereits dem logischen Konstrukt einer präklinischen Studie und wird in Abb. 8 gezeigt.

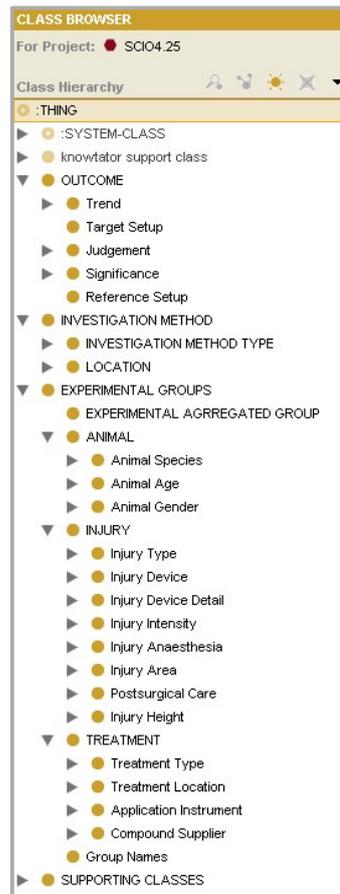


Abb. 8 **Darstellung der Klassenhierarchie mit Protégé.** Die Hauptklassen *Outcome*, *Investigation Method*, *Experimental Groups* und *Supporting Classes* bilden vier, voneinander unabhängige und gleichwertige Hauptklassen mit ihren dazugehörigen Unterklassen.

In Abb. 8 ist eine Übersicht der SCIO-Ontologie mit ihren Haupt- und Unterklassen aufgeführt. Ein mit seiner Spitze nach rechts zeigendes Dreieck indiziert, dass einer Klasse weitere Klassen untergeordnet sind, aber nicht angezeigt werden. Die Klasse ist „geschlossen“. Zeigt die Spitze eines Dreiecks nach unten werden die Unterklassen einer Hauptklasse sichtbar. Die entsprechende Klasse ist „geöffnet“. Von oben nach unten betrachtet zeigen sich die vier voneinander unabhängigen und gleichwertigen Hauptklassen. Allen voran steht die Klasse *Outcome*, die dem logischen Aufbau einer präklinischen Studie entsprechend das Zentrum der Ontologie bildet, dass wiederum mit der unter ihr aufgeführten Klasse *Investigation Method* die Klasse *Experimental Groups* untersucht. Die Klasse *Supporting Classes* entspricht einer Hauptklasse, die eine Vielzahl an Unterklassen aufweist, die Auskünfte über Details, wie Maßeinheiten, respektive Zahlenwerte geben. Diese sind von Studie zu Studie unterschiedlich und werden den experimentellen Gruppen und Ergebnissen zugeordnet.

2.3.2 Die Hauptklasse Animal

Beginnend mit der Klasse *Animal* wurde eine Klassenhierarchie bestehend aus Hauptklassen und Unterklassen definiert, ein Prozess, der bei allen weiteren Klassen angewendet wurde.

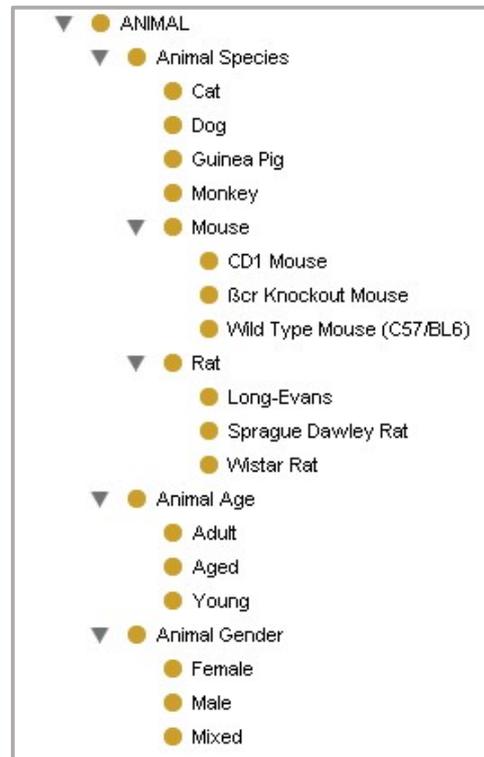


Abb. 9: **Die Klassenhierarchie der Hauptklasse *Animal*.** Aus ihr gehen die drei Unterklassen *Animal Species*, *Animal Age* und *Animal Gender* hervor. Neben den Klassen zeigen sich graue, mit ihren Spitzen nach unten gerichtete Dreiecke, die eine „geöffnete Klasse“ mit weiteren Unterklassen anzeigen.

Die in Abb. 9 gezeigte *Animal*-Klasse ist die übersichtlichste aller Klassen, die bei der Beschreibung einer präklinischen Studie mit ihren Konzepten zur Gesamtrepräsentation des Domänenwissens beitragen. Bei der Analyse des vorgegebenen Textkorpus stellte sich heraus, dass in den unterschiedlichen präklinischen Studien eine übersichtliche Anzahl an unterschiedlichen Tierspezies eingesetzt wurden. Bei der Annotation von Texten sollte stets mit der Klasse auf höchster Detailebene annotiert werden. Sollte im Fließtext einer wissenschaftlichen Publikation das Geschlecht eines Tieres annotiert werden, würde dies nicht mit der Hauptklasse *Animal Gender* geschehen, sondern mit einer der zutreffenden Unterklassen *Male*, *Female* oder *Mixed*. Dieses Prinzip erstreckt sich auf die Anwendung aller Klassen der SCIO-Ontologie und dient einer manuellen

Datenextraktion von höchst aufschlussreichen Konzepten, um eine genaue Repräsentation von Studien zu ermöglichen.

2.3.4 Die Hauptklasse Injury

Eine weitere, durchaus umfangreichere Hauptklasse ist die Klasse *Injury*. Sie beschreibt die Materialien und Methoden, die bei der Verletzung von experimentellen Tiergruppen ihre Anwendung finden.

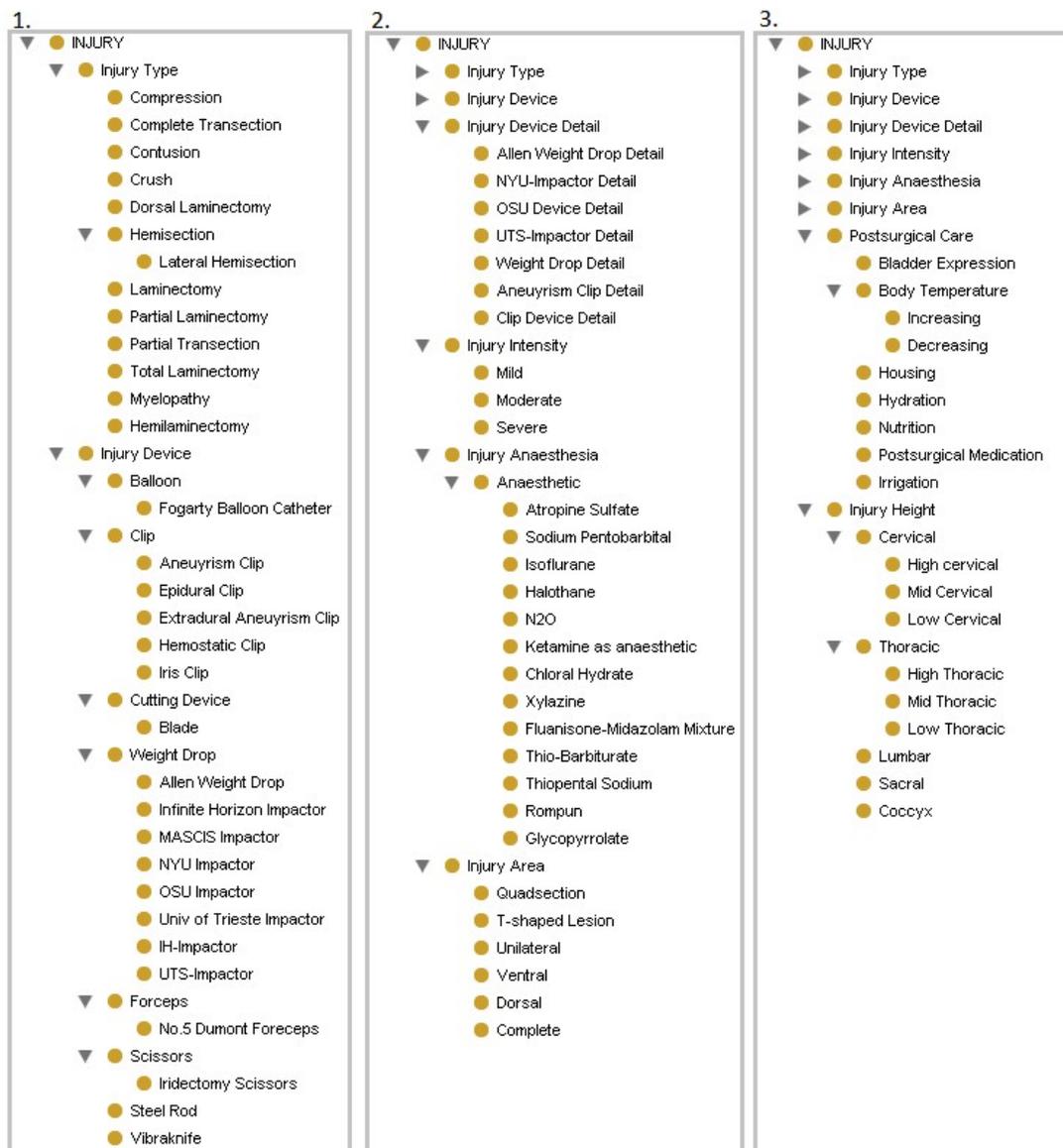


Abb. 10: **Klassenhierarchie der Hauptklasse Injury.** Die Klasse Injury wird in ihrer Reihenfolge von oben nach unten und von links nach rechts in drei neben einander stehenden Partitionen gezeigt. Sie besteht aus den acht Hauptklassen Injury Type, Injury Device, Injury Device Detail, Injury Intensity, Injury Anaesthesia, Injury Area, Postsurgical Care, Injury Height mit ihren dazugehörigen Unterklassen.

Nachdem wie in Abb. 10 dargestellt die Hauptkonzepte einer präklinischen Studie über TRV bestimmt wurden, erfolgte die Modellierung der Ontologie. Zunächst wurde das Oberklassen- und Unterklassengefüge zur Bestimmung des Wirkgrades der Ontologie definiert. Hier sollte bereits zu Beginn der Wissensumfang abgesteckt werden, um eventuelle Wissenslücken zu füllen und redundantes Wissen zu vermeiden. Die Zuordnung der Unterklassen zu ihren komplementären Oberklassen erfolgte basierend auf ihren inhaltlichen und sinngemäßen Zugehörigkeiten, als Teile von Oberklassen (*a part of*). Diese Zuordnungen erfolgten bis ein essenzielles Minimum an notwendigem Domänenwissen, durch die Oberklassen und Unterklassen vollständig repräsentiert werden konnten. Dies ließ sich feststellen, wenn während der manuellen Datenextraktion mit einer vorläufigen Version der Ontologie festgestellt wurde, dass bestimmte Konzepte nicht annotiert werden konnten. In solchen Fällen wurden entsprechende Klassen in die Ontologie integriert. Die Anordnung der Klassen stellt den primären Aufbau der Ontologie dar, während die später hinzugefügten Eigenschaften von Klassen den sekundären Aufbau der Ontologie, respektive das Relationsgefüge unter den Klassen herstellten. Neben der hierarchischen Kategorisierung von Klassen stellte die Reihenfolge bei der Erwähnung von Konzepten in wissenschaftlichen Veröffentlichungen ein wichtiges Kriterium bei der Ontologieentwicklung dar. Ihre Berücksichtigung sollte die Anwendung der Ontologie bei der Annotation von Texten erleichtern. In diesem Zusammenhang ist zu beachten, dass die unterschiedlichen Konzepte einer Hauptklasse bei Ihrer Erwähnung im Text keiner vorbestimmten Reihenfolge unterliegen, jedoch lassen sich wiederholende Muster bei der Beschreibung des experimentellen Aufbaus und der Ergebnisse in unterschiedlichen Publikationen feststellen. Bei der Beschreibung der Klasse *Animal* werden alle relevanten Unterklassen, wie die Spezies, das Geschlecht, Alter und Gewicht konzis mit ein bis drei zusammenhängenden Sätzen erläutert. Dies ist sicherlich auf die überschaubare Zusammensetzung der *Animal*-Klasse zurückzuführen. Die Inkorporation einer Reihenfolge von Konzepten stellte während der Konstruktion der Ontologie bei größeren Hauptklassen, wie bei der *Injury*-Klasse, eine Herausforderung dar. Aufgrund der vielen Unterklassen ist eine uniforme Reihenfolge bei der Erwähnung von Konzepten nicht immer festzustellen. Die Art der induzierten Verletzung und der betroffene anatomische Abschnitt des Rückenmarks finden in den meisten Fällen als erstes ihre Erwähnung, gefolgt von den verwendeten Instrumenten oder Apparaturen, die hierfür eingesetzt werden. Weitere Details zu den Instrumenten, wie

Geräteeinstellungen und die applizierte Intensität bei der experimentell verursachten Verletzung des Rückenmarks werden oft anschließend beschrieben. Obwohl die verabreichte Anästhesie ein Prärequisit für die Ausführung einer Verletzung des Rückenmarks ist, werden Angaben hierzu in den meisten Fällen erst nach der Beschreibung des Verletzungsmodells gemacht. Die postoperative Versorgung der Tiere wird im Allgemeinen oft erst zum Schluss beschrieben. Hierzu gehören die Unterbringung und das Füttern der Tiere. Zusätzlich wird in diesem Abschnitt die postoperative Verabreichung von Medikamenten erwähnt. Ein individuelles Merkmal der Klasse *Injury* ist das Ausmaß ihrer Beschreibung, das im Vergleich zu den Klassen *Animal* und *Treatment* in wissenschaftlichen Veröffentlichungen umfangreicher ist, da in diesem Abschnitt der *Experimental Group* der Versuchsaufbau ausführlich beschrieben wird. Die Klasse *Injury* wird in Publikationen unmittelbar nach der Beschreibung der Klasse *Animal* und vor der Erläuterung der angewendeten Therapien präsentiert.

2.3.4 Die Hauptklasse Treatment

Die Klasse *Treatment* ist die Hauptklasse mit den meisten Unterklassen und bildet in wissenschaftlichen Publikationen nach vorausgehender Beschreibung der verwendeten Tierspezies und den experimentell verursachten Verletzungen den Abschluss des experimentellen Ablaufs. Die Evaluation der unterschiedlichen Behandlungsmethoden erfolgt im Ergebnisteil einer wissenschaftlichen Publikation unter Anwendung bestimmter Untersuchungsmethoden, die Aufschluss darüber geben, ob eine medikamentöse Therapie im Vergleich zu einer anderen erfolgreicher war und funktionale Defizite beseitigt werden konnten oder ob auf histologischer Ebene nachvollziehbare Verletzungen wieder verheilen konnten. Die Klasse *Treatment* bildet den Abschluss des experimentellen Aufbaus und schließt die Beschreibung der experimentellen Gruppen ab. Aufgrund ihrer Größe und zur übersichtlichen Darstellung wird sie, wie die Klasse *Injury*, in mehreren Partitionen in Abb. 11 dargestellt. Die bei der medikamentösen Behandlung verwendeten Substanzen bilden mit Abstand die umfangreichste Unterklasse (*Substance*) der Hauptklasse *Treatment*. In einer präklinischen Studie werden in der Regel mehrere Substanzen bei unterschiedlichen experimentellen Gruppen gleichzeitig oder zu verschiedenen Zeitpunkten verabreicht.

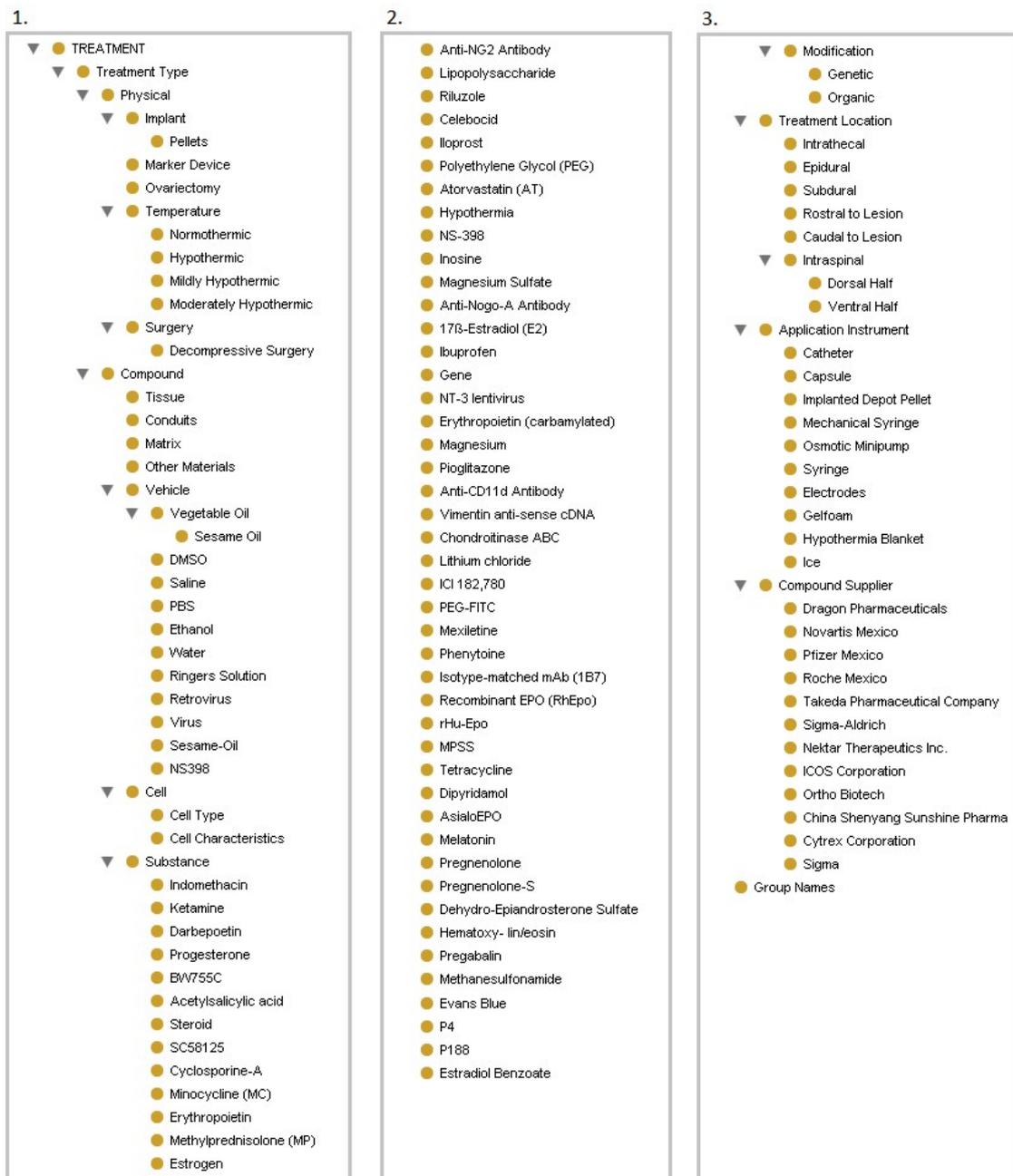


Abb 11. **Klassenhierarchie der Hauptklasse Treatment.** Die in der Protégé-Software von oben nach unten angezeigte Reihenfolge der *Treatment*-Klasse ist in dieser Abbildung von oben nach unten und von links nach rechts in drei Partitionen dargestellt. Sie besteht aus den vier Hauptklassen *Treatment Type*, *Treatment Location*, *Application Instrument*, *Compound Supplier* und den dazugehörigen Unterklassen.

Die in Abb. 11 gezeigte *Treatment*-Klasse beschreibt mit ihren Oberklassen und Unterklassen den allgemeinen Behandlungsablauf, die speziellen Therapieformen (*Treatment Type*) und die in diesem Zusammenhang verwendeten medizinischen Präparate (*Compound*), die bei unterschiedlichen experimentellen Gruppen angewendet werden. Des Weiteren beschreiben die anatomische Region (*Treatment Location*) der

medikamentösen Verabreichung und die dabei verwendeten Instrumente (*Application Instrument*) neben den pharmazeutischen Unternehmen (Compound Supplier), die ihre Präparate zur Verfügung stellen, weitere Aspekte des Therapiemodells. Im Fließtext von wissenschaftlichen Publikationen erfolgt die Beschreibung der applizierten Therapien für jede experimentelle Gruppe in wenigen, konzisen Sätzen und schließt die Ausführung des experimentellen Ablaufs und Aufbaus ab.

Eine weitere zur *Experimental Group* gehörende Unterklasse ist die Klasse *Group Names*. Da experimentelle Gruppen in verschiedenen Abschnitten einer wissenschaftlichen Publikation unterschiedlich benannt werden, wurde diese Klasse eingeführt, um die heterogenen Benennungen den entsprechenden experimentellen Gruppen zuordnen zu können. Demnach kann eine mit Kochsalz behandelte experimentelle Kontrollgruppe im Material- und Methodenteil mit dem Namen „*saline-group*“ beschrieben werden, während sie im Ergebnisteil beim Vergleich zwischen Zielgruppen und Referenzgruppen mit dem Namen „*animals treated with saline solution*“ erwähnt wird. In diesem Fall würde der ursprüngliche Name im Material- und Methodenteil mit der Klasse *Experimental Group* annotiert werden. Die vom ursprünglichen Namen abweichende Erwähnung im Ergebnisteil würde mit der Klasse *Group Names* annotiert werden, damit sie der entsprechenden experimentellen Gruppe zugeordnet werden kann. Mithilfe der Klasse *Group Names* sollen die verschiedenen Namensgebungen zum Zwecke einer sachgerechten Zuordnung und adäquaten Übersicht der experimentellen Gruppen standardisiert werden.

2.3.5 Die komplementären Klassen

Das grundlegende Ziel der SCIO-Ontologie, Wissenschaftlern einen genauen Überblick über eine präklinische Studie zu gewährleisten, indem studienrelevante Daten adäquat und übersichtlich präsentiert werden, um die Entscheidung einer potentiellen Fortführung in einer klinischen Studie zu erleichtern, barg die Herausforderung einen angemessenen Detailgrad bei der Präsentation von studienrelevanten Daten anzuwenden. Bei der Analyse der 140 wissenschaftlichen Publikationen zeigten sich Datenpunkte, die aufgrund ihres hohen Detailgrades und ihrer Verschiedenheit Inkonsistenzen bei der Annotation von wissenschaftlichen Texten hervorrufen können. Bei diesen Datenpunkten handelt es sich um Konzepte, die ein höheres Maß an Informationen beinhalten und den gewünschten Detailgrad der SCIO-Ontologie übersteigen. Hierzu gehören Konzepte, die Zahlenwerte mit entsprechenden

Maßeinheiten repräsentieren und von Publikation zu Publikation unterschiedlich sind und keiner einfachen Definition und Klassifizierung unterzogen werden können. Zu diesen Konzepten gehören der p-Wert, genaue Angaben zum Körpergewicht, die unterschiedlichen Beobachtungszeitpunkte, die Anzahl N von Tieren in einer experimentellen Gruppe, Angaben zu Geräteeinstellungen von experimentellen Apparaturen, sowie genaue Angaben zu Abschnitten des untersuchten Rückenmarks. Darüber hinaus wurden die Verabreichungsformen von Medikamenten, z.B. die orale, intravenöse und subkutane Verabreichung den *Supporting Classes* untergeordnet. Diese spezifischen Angaben beinhalten Informationen, die für den experimentellen Hergang und für die Ermittlung von Ergebnissen durchaus notwendig, aber für einen verständlichen Überblick über eine Studie zu spezifisch sind. Im Hinblick auf die Implementierung einer automatischen Datenextraktion sollte die SCIO-Ontologie in ihrer Funktion als Rückgratstruktur diese Datenpunkte sammeln, um sie im weiteren Verlauf bei der automatischen Datenextraktion anwenden zu können. Die Klasse *Supporting Classes* ermöglicht bei ihrer Anwendung die Annotation und Extraktion von äußerst spezifischen Daten, die komplementär zum Oberklassen- und Unterklassengefüge der Ontologie eine hohe Granularität aufweisen, um Wissenslücken zu vermeiden. Ein wichtiges Beispiel ist die Unterklasse *Application Timespoints*, die die genauen Zeitpunkte der Verabreichung von Medikamenten anzeigt. Die zuvor erwähnten Beobachtungszeitpunkte, die ein wichtiger Bestandteil der Untersuchungsmethoden sind, werden ebenfalls durch die *Supporting Classes* repräsentiert. Während die unterschiedlichen Abschnitte des Rückenmarks, wie thorakal, lumbal und sakral in der Unterklasse *Injury Height* der Hauptklasse Injury beschrieben werden, werden die genauen Angaben eines Rückenmarkabschnittes, z.B. der Abschnitt TH 9-12 des thorakalen Rückenmarks, durch die entsprechende Unterklasse *Injury Height Detail* der *Supporting Classes* manuell aus den Texten extrahiert. Bei jeder Hauptklasse der SCIO ist die Integration von Werten aus den *Supporting Classes* vorgesehen, um eine vollständige und zugleich individuelle Zusammenfassung jeder präklinischen Studie zu ermöglichen. In Abb. 12 ist eine Übersicht der Hauptklasse *Supporting Classes* dargestellt. Von oben nach unten betrachtet zeigen sich alle Klassen, die ihre Anwendung bei der Komplementierung der entsprechenden Hauptklassen finden. Langfristig betrachtet sollen die mit den *Supporting Classes* annotierten Werte für prospektive und erweiterbare Anwendungen der SCIO unter Umständen benötigte Datenpunkte zur Verfügung stellen.

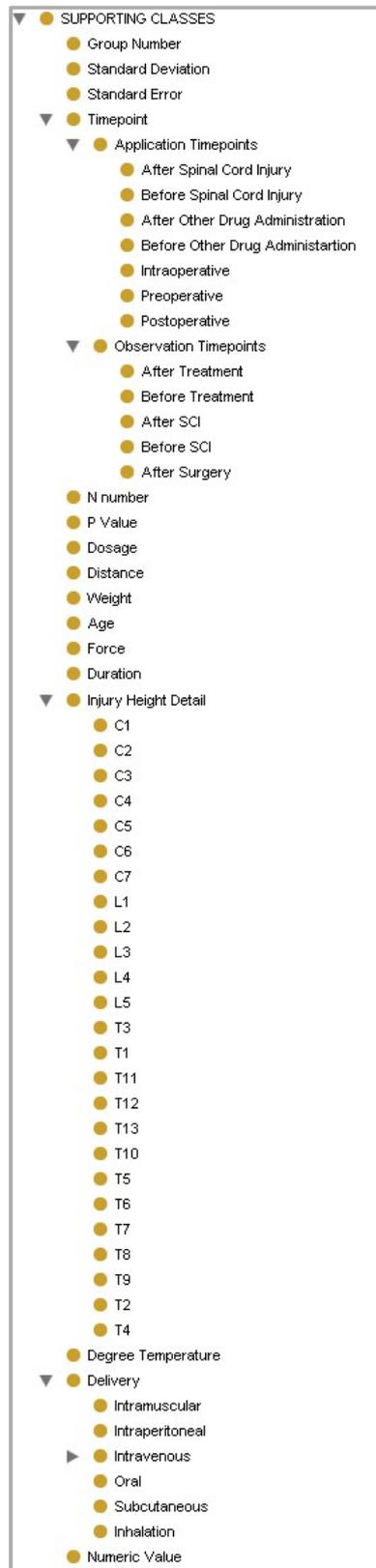


Abb. 12 **Klassenhierarchie der *Supporting Classes***. Spezifische und individuelle Informationen in Form von Zahlenwerten, wie das Alter von Tieren, Zeitpunkte, verabreichte Dosen, Gewicht und Temperaturwerte, die nicht unmittelbar zur Domäne der traumatischen Rückenmarksverletzungen gehören, werden mit einer entsprechenden Unterklasse der *Supporting Classes* annotiert.

2.3.6 Die Hauptklasse Investigation Method

In wissenschaftlichen Publikationen werden nach der Beschreibung des experimentellen Setups die Effektivität der applizierten Therapien durch investigative Untersuchungsmethoden bewertet und validiert. Sie geben Aufschluss über positive oder ggf. auch negative Veränderungen in Bezug auf die Regeneration des Rückenmarks von Versuchstieren, die nach experimentell induzierten traumatischen Rückenmarksverletzungen einer Therapie unterzogen wurden. Darüber hinaus ermöglichen sie einen direkten Vergleich von unterschiedlichen experimentellen Gruppen. Die Untersuchungsmethoden werden durch die Hauptklasse *Investigation Method* mit ihren dazugehörigen Unterklassen repräsentiert. Sie besteht aus zwei Unterklassen, *Investigation Type* und *Location*. Wie in Abb. 7 gezeigt, bildet die Hauptklasse *Investigation Method* den Angelpunkt zwischen dem experimentellen Setup und den Ergebnissen und hat somit eine Schlüsselfunktion bei der Schaffung einer zusammenhängenden Übersicht aller Datenpunkte des experimentellen Aufbaus und Ablaufs einschließlich aller relevanten Ergebnisse, die zur Beurteilung der Erfolgsaussichten bei der Anwendung bestimmter Therapien benötigt werden. Sie kann als autonome Entität einer präklinischen Studie über traumatische Rückenmarkverletzungen angesehen werden, da sie weder als Unterklasse der Hauptklasse *Experimental Groups* noch als Unterklasse der Hauptklasse *Outcome* fungiert. Die Untersuchungsmethoden werden in funktionale (*Functional*) und nicht funktionale (*Non-Functional*) Untersuchungsmethoden aufgeteilt. Zu den funktionalen Tests gehören elektrophysiologische und physiologische Untersuchungen. Lokomotorische Untersuchungen, die das Bewegungsverhalten von Tieren nach experimentell verursachten Traumata mit konsekutiven Behandlungen beobachten, werden in den meisten präklinischen Studien zur Beurteilung des Therapieerfolges herangezogen. Weitere funktionale Untersuchungsmethoden sind sowohl reflektorische und sensorische Tests als auch die Applikation neurologischer Skalen zur Beurteilung eines Therapieerfolges. Genaue Angaben zu den spezifischen anatomischen Regionen, die histologisch untersucht wurden, werden durch die Hauptklasse *Location* vermittelt. Hierbei handelt es sich um die genauen Einwirkungsorte der traumatischen Verletzung auf das Rückenmark. Zu den nicht funktionalen Untersuchungsmethoden zählen histologische Untersuchungen, die z.B. Aufschluss über neuronale Veränderungen, wie die axonale Regeneration geben. Weitere Veränderungen, die nicht neurologischen

Ursprungs sind (*Non-Functional*) können ebenfalls auf histologischer Ebene festgestellt werden. Die Beobachtung der sekundären Degeneration nach traumatischen Einwirkungen auf das Rückenmark beschreibt reaktive Prozesse auf histologischer Ebene, wie die Apoptose, Blutungen und den oxidativen Stress. Eine weitere Unterklasse der Oberklasse *Non-Functional* sind die molekularen Untersuchungsmethoden.

Die Hauptklasse *Investigation Method* relativiert die Beobachtungen, die aus dem experimentellen Ablauf und Aufbau gewonnen werden können und erlaubt zugleich eine detaillierte Auswertung der erworbenen Erkenntnisse im Ergebnisteil einer wissenschaftlichen Publikation. Im Umkehrschluss bestimmt die Hauptklasse *Investigation Method* das Study Design einer präklinischen Studie dahingehend, dass das verwendete Verletzungsmodell und die applizierten Therapien durch die vorgesehenen Untersuchungsmethoden beurteilt werden können. Die *Investigation Method* findet ihre Erwähnung im Text einer Publikation nach der Beschreibung des experimentellen Ablaufs und vor der Präsentation der Ergebnisse. Zusätzlich wird sie im Ergebnisteil einer Veröffentlichung beim Vergleich von unterschiedlich behandelten experimentellen Gruppen erwähnt. Sie ist neben den unterschiedlichen experimentellen Tiergruppen einer der maßgeblichen und limitierenden Faktoren in Bezug auf die Anzahl der Ergebnisse, die aus einer präklinischen Studie hervorgehen können. Jeder Vergleich einer experimentellen Zielgruppe mit einer experimentellen Referenzgruppe erfolgt auf der Basis einer bestimmten Untersuchungsmethode, die die Grundlage des Vergleichs darstellt. Obwohl die Untersuchungsmethoden den experimentellen Abschnitt einer Publikation mit den Ergebnissen verbinden, haben sie eine größere Bedeutung bei der Darstellung der Ergebnisse, da jedes Ergebnis auf einer Untersuchungsmethode aufbaut und sie somit als eine Eigenschaft von Ergebnissen betrachtet werden kann. Zusätzlich ist die Signifikanz eines Ergebnisses auf die applizierte Untersuchungsmethode zurückzuführen. Eine präklinische Studie kann durch unterschiedliche investigative Verfahren beurteilt werden, die bei steigender Anzahl eine proportionale Zunahme der Ergebnisse zur Folge haben. Zusammenfassend kann konstatiert werden, dass die Hauptklasse *Investigation Method* einen essenziellen Parameter bei der Präsentation der Ergebnisse darstellt, die Qualität der verwendeten Therapien beurteilt und für die Quantität der Ergebnisse verantwortlich ist. Abb. 13 zeigt die Hauptklasse *Investigation Method*.

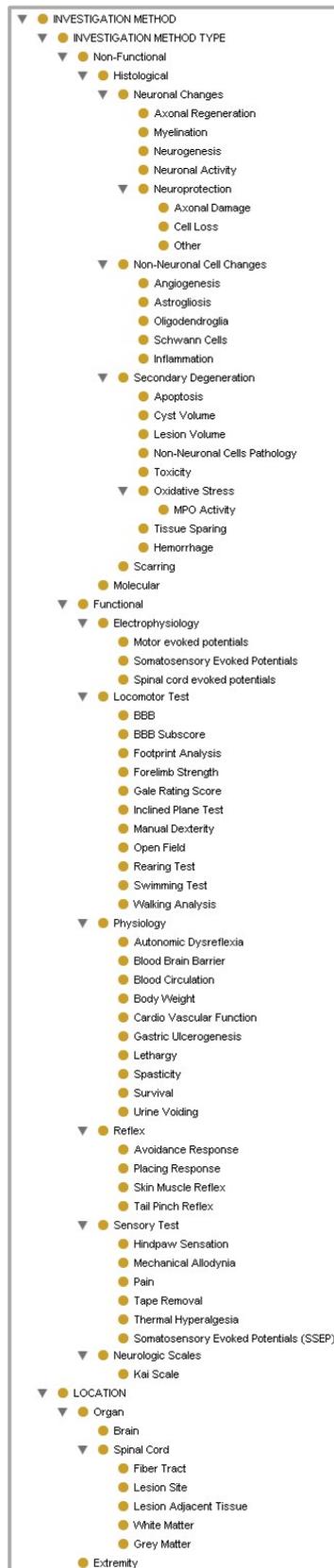


Abb. 13 **Klassenhierarchie der Hauptklasse Investigation Method.** Sie repräsentiert die Untersuchungsmethoden, die zur Auswertung der verschiedenen Therapieansätze herangezogen werden und beim Vergleich zwischen unterschiedlichen experimentellen Tiergruppen den Referenzwert bilden.

2.3.7 Die Hauptklasse Outcome

Nachdem der experimentelle Aufbau beschrieben wurde und die Untersuchungsmethoden zur Auswertung der gewonnenen Daten appliziert wurden erfolgt eine systematische Darstellung der Ergebnisse. Die Hauptklasse *Outcome* beschreibt ein Ergebnis, indem verschiedene experimentelle Gruppen, bestehend aus den Hauptklassen *Animal*, *Injury* und *Treatment* miteinander verglichen werden. Sie entspricht einer Zusammenfassung aller Hauptklassen der experimentellen Gruppen, die basierend auf einem Referenzwert, der *Investigation Method*, miteinander verglichen werden. Die Ergebnisse bilden somit das Zentrum der SCIO, da hier die Kernaussage einer wissenschaftlichen Publikation zusammengefasst dargestellt wird und Rückschlüsse auf den experimentellen Aufbau und Ablauf gezogen werden können. Abb. 14 zeigt den Aufbau der Hauptklasse *Outcome*.

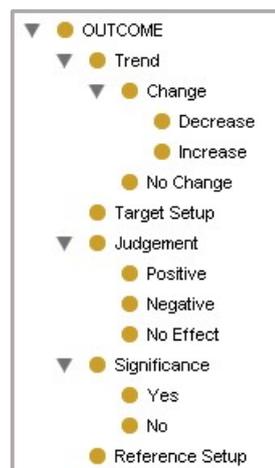


Abb. 14 **Klassenhierarchie der Hauptklasse Outcome.** Alle notwendigen Konzepte für den Vergleich von unterschiedlichen experimentellen Gruppen werden durch die Hauptklasse *Outcome* bereitgestellt.

Wie in Abb. 14 von oben nach unten dargestellt besteht die Klasse *Outcome* aus den Unterklassen *Trend*, *Target Setup*, *Judgement*, *Significance* und *Reference Setup*. Die Unterklasse *Trend* beschreibt eine Veränderung, die mit einer Untersuchungsmethode festgestellt wurde. In diesem Zusammenhang kann im Rahmen einer Untersuchung der Lokomotion von Versuchstieren nach vorausgegangener Verletzung des Rückenmarks mit anschließender Therapie ein Anstieg (*Increase*), eine Verminderung (*Decrease*) oder keine Veränderung (*No Change*) der Mobilität durch die Klasse *Trend* festgestellt

werden. Die Unterklasse *Target Setup* repräsentiert die Konfiguration einer bestimmten Zielgruppe, die mit der Konfiguration einer bestimmten Referenzgruppe (*Reference Setup*) verglichen wird. Wie bereits zuvor erwähnt entspricht ein Vergleich zwischen unterschiedlichen experimentellen Gruppen einem Ergebnis (*Outcome*). Die Unterklasse *Judgement* beschreibt eine subjektive Beurteilung, die durch den Autor in Bezug auf ein Ergebnis gemacht wird. Die Klasse *Significance* indiziert, ob bei einem vorliegenden P-Wert ein Ergebnis als signifikant oder nicht signifikant eingestuft werden kann. Die Hauptklasse *Outcome* fasst nach vollständiger Annotation einer wissenschaftlichen Publikation alle essenziellen Konzepte für die Beurteilung einer präklinischen Studie zusammen. Die annotierten Konzepte der Hauptklassen *Experimental Groups* und *Investigation Method* fließen in das Wissenskonstrukt der Hauptklasse *Outcome* ein und gewährleisten einen detaillierten Überblick einer präklinischen Studie.

2.4 Attribute und das Relationsgefüge der SCIO

Die Klassenhierarchien aller individuellen Hauptklassen entsprechen dem primären, vertikalen Aufbau der Ontologie. Sie repräsentieren das notwendige Maß an Konzepten, dass für eine adäquate Repräsentation des Domänenwissens über traumatische Rückenmarksverletzungen aus präklinischen Studien benötigt wird. Im primären Aufbau werden die Hauptklassen mit ihren Unterklassen unabhängig voneinander dargestellt. Würden Annotationen von wissenschaftlichen Texten ausschließlich mit dem primären Aufbau der Ontologie erfolgen, hätte dies zur Folge, dass das extrahierte Wissen nicht zusammenhängend vorliegen würde. Ergebnisse würden wichtige Entitäten wie die experimentellen Gruppen und Untersuchungsmethoden unabhängig voneinander präsentieren und Beurteilungen von Therapieansätzen wären aufgrund der fehlenden Zusammenhänge nicht möglich. Die primäre Organisationsstufe und die fehlenden Bezüge zwischen den unterschiedlichen Hauptklassen mitsamt ihren Unterklassen ergaben die Notwendigkeit für die Entwicklung einer Strategie zur Verknüpfung des extrahierten Wissens. Hierfür eignete sich die Inkorporation von Eigenschaften in das Klassengefüge der SCIO. Eigenschaften, auch Attribute oder *slots* genannt, ermöglichen die Festlegung von Angelpunkten zur Herstellung einer Verbindung zwischen den unterschiedlichen Klassen, indem sie sogenannten Schlüsselklassen zugeordnet werden. Bei den Schlüsselklassen handelt es sich um

Unterklassen, die bei der Annotation von wissenschaftlichen Texten regelmäßig angewendet werden und für die Erfassung des experimentellen Aufbaus und Ablaufs ein Prerequisite sind. So stellt die Tierspezies, repräsentiert durch die Unterklasse *Animal Species*, eine Mindestanforderung dar, die benötigt wird, um das Tiermodell einer präklinischen Studie adäquat zu beschreiben. Die Integration von Attributen in die Struktur der Ontologie erfolgt nicht im Sinne einer hierarchischen Unterordnung an eine Hauptklasse, sondern als Definition einer zusätzlichen Wissensseinheit, die nicht in der Klassenhierarchie angezeigt wird, aber parallel zu ihr existiert. Die Attribute einer Hauptklasse sind Kopien ihrer respektiven Unterklassen, die nach ihrer Annotation aggregiert werden und einzelnen Schlüsselklassen als Attribute zugeordnet werden. Hieraus entsteht die sekundäre Organisationsstruktur der SCIO. Die zunächst auf dem Niveau der individuellen Schlüsselklassen aggregierten Attribute können im weiteren Verlauf den Hauptklassen *Experimental Groups* und *Outcome* zugeordnet werden. Dies ermöglicht eine horizontale Vernetzung der vertikalen Wissensstruktur. Das aggregierte Wissen kann den einzelnen Ergebnissen beim Vergleich der unterschiedlichen experimentellen Gruppen zugeordnet werden. Somit geht aus einem Ergebnis hervor, aus welchen Tierspezies die experimentellen Gruppen bestehen und welche experimentell verursachten Verletzungen und Therapien Teil des experimentellen Aufbaus sind. Abb. 15 zeigt die Attribute der Hauptklasse *Animal*.

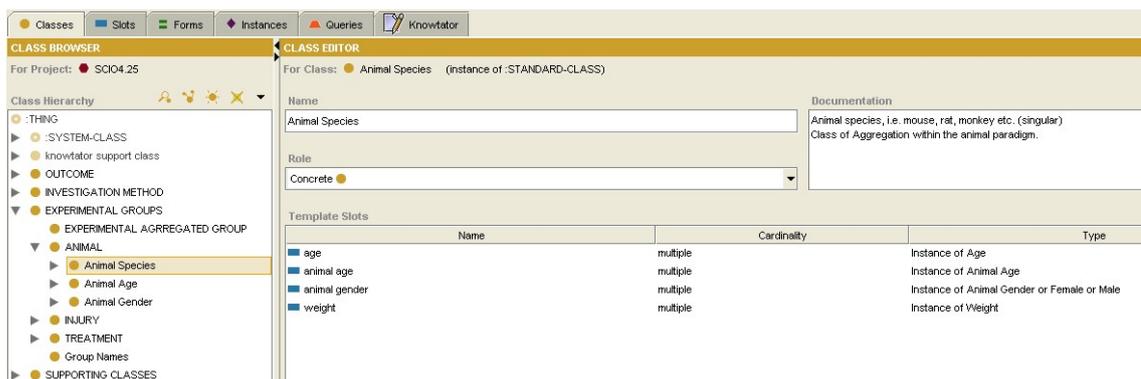


Abb. 15 Die Attribute der Klasse *Animal*. Die auf der linken Seite im *Class Browser* hervorgehobene Unterklasse *Animal Species* der Hauptklasse *Animal* ist die Schlüsselklasse, der die Attribute *age*, *animal age*, *animal gender* und *weight* auf der rechten Seite in der Rubrik *Template Slots* zugeordnet werden.

Nachdem die Unterklassen *age*, *animal age*, *animal gender* und *weight* im Fließtext einer wissenschaftlichen Publikation mit den entsprechenden Unterklassen der

Hauptklassen *Animal* und *Supporting Classes* annotiert wurden, können sie in aggregierter Form der annotierten Spezies respektive der Schlüsselklasse, wie in Abb. 15 gezeigt, zugeordnet werden. Dieses Verfahren findet bei allen Hauptklassen seine Anwendung.

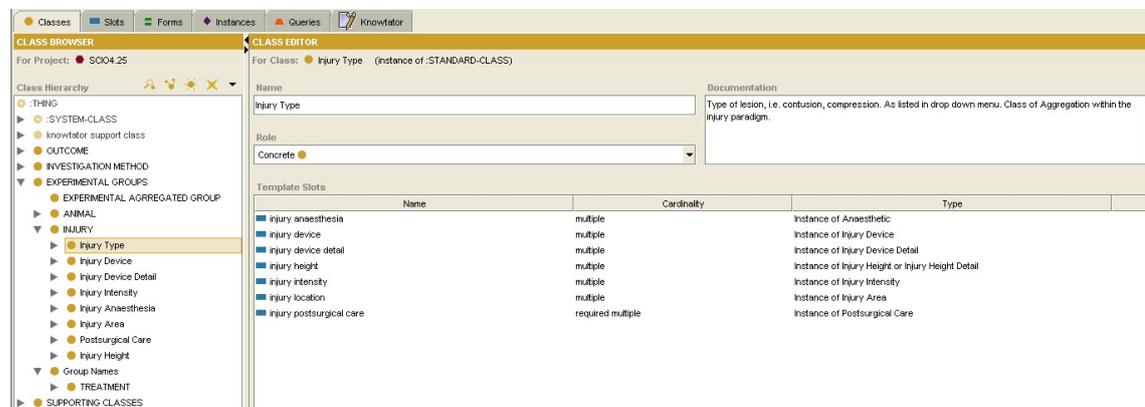


Abb. 16 Die Attribute der Klasse *Injury*. Die Schlüsselklasse der Hauptklasse *Injury* ist die hervorgehobene Unterklasse *Injury Type*, die bei der Beschreibung des experimentellen Ablaufs und Aufbaus ein Prärequisit ist. Ihr werden die Attribute aus der Rubrik *Template Slots* zugeordnet, so dass die Information in aggregierter Form vorliegt.

Wie bei den Attributen der Hauptklasse *Animal* handelt es sich bei den in Abb. 16 gezeigten Attributen der Hauptklasse *Injury* um Kopien ihrer respektiven Unterklassen, die nach der Annotation mit den entsprechenden Unterklassen einer Schlüsselklasse zugeordnet werden. Die aggregierten Attribute werden bei der Verwendung der Schlüsselklasse angezeigt. Das gleiche Konzept der Wissensaggregation wird in Abb. 17 für die Hauptklasse *Treatment* dargestellt.

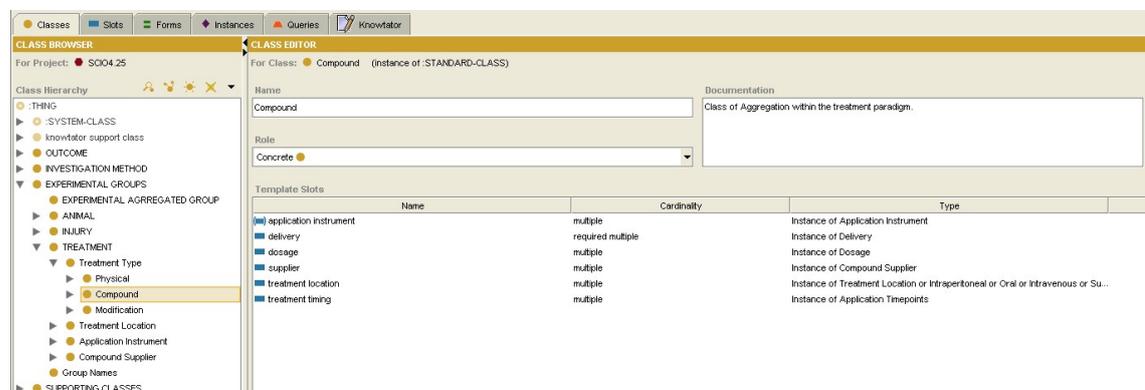


Abb. 17 Die Attribute der Klasse *Treatment*. Die Schlüsselklasse ist die Unterklasse *Compound*, die für eine adäquate Repräsentation des Behandlungsmodells unverzichtbar ist.

Die SCIO speichert und repräsentiert das extrahierte Domänenwissen primär durch den vertikalen Aufbau der Klassenhierarchien und sekundär durch die horizontale Quervernetzung mithilfe von Attributen. Die Aggregation von Attributen beginnt auf dem Niveau der Unterklassen und wird auf den höheren Ebenen der Oberklassen in einem additiven Verfahren fortgesetzt, indem die aggregierten Attribute der *low level classes* (*Animal, Injury, Treatment* und *Investigation Method*) an die Attribute der *upper level classes* (*Experimental Groups* und *Outcome*) angehängt werden. Diese Anordnung des akkumulierten Wissens ähnelt einem zwiebelschalenartigen Aufbau. Das auf dem Niveau von Unterklassen aggregierte Wissen wird auf der Ebene der Oberklassen akkumuliert, so dass das von innen (Unterklassen) nach außen (Oberklassen) in seiner Quantität immer größer werdende Wissen in Schalen und definierten Wissenseinheiten vorliegt. Mit den größer werdenden Schalen nimmt die Quantität des extrahierten Wissens auf den höheren Ebenen zu und wird gleichzeitig in geordneten Wissenseinheiten abgespeichert. Somit kann eine hohe Quantität extrahierten Wissens erfasst, kategorisiert und gezielt abgerufen werden. Die vermeintliche Informationsflut wird bei dieser Herangehensweise der Wissensmodellierung kontrolliert und ihre Zugänglichkeit vereinfacht. Diese Form der Aggregation erlaubt eine Zuordnung der extrahierten Konzepte über das Tier-, Verletzungs- und Therapiemodell auf bestimmte experimentelle Gruppen.

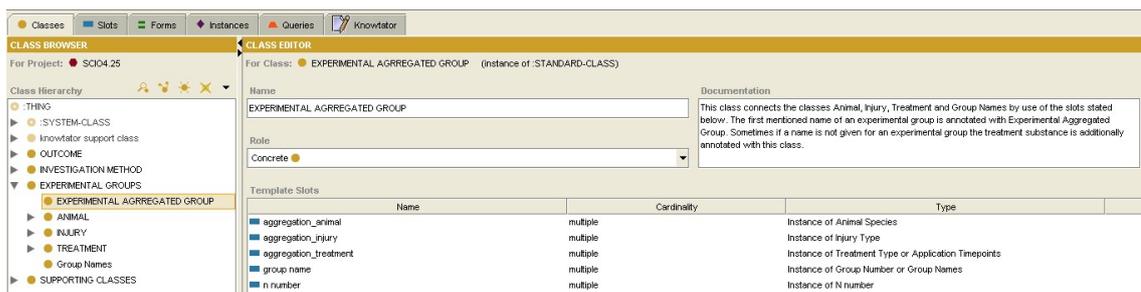


Abb. 18 Die Attribute der Klasse *Experimental Groups*. Zur Unterklasse *Experimental Aggregated Group* gehören die in der Rubrik Template Slots gezeigten Attribute: *aggregation animal, aggregation injury, aggregation treatment, group name* und *n number*. Diese Attribute fassen das extrahierte Wissen der entsprechenden Hauptklassen zusammen und ordnen es bestimmten experimentellen Gruppen zu. Somit werden die individuellen experimentellen Gruppen genau beschrieben und voneinander differenziert.

Abb. 18 zeigt, wie die Klasse *Experimental Aggregated Group* das Wissen des Material- und Methodenteils einer wissenschaftlichen Publikation bündelt. Alle

Informationen, die für das Verständnis des experimentellen Aufbaus und Ablaufs einer präklinischen Studie notwendig sind, gehen aus ihr hervor.

Nachdem die experimentellen Rahmenbedingungen und Abläufe durch die experimentellen Gruppen mitsamt ihren Attributen dargestellt werden, können durch die Applikation bestimmter investigativer Verfahren die unterschiedlichen Therapieansätze evaluiert und durch die Oberklasse *Outcome* präsentiert werden. Hierbei werden Attribute des experimentellen Abschnitts einem Ergebnis, oder *Outcome*, mittels Aggregation zugeordnet.

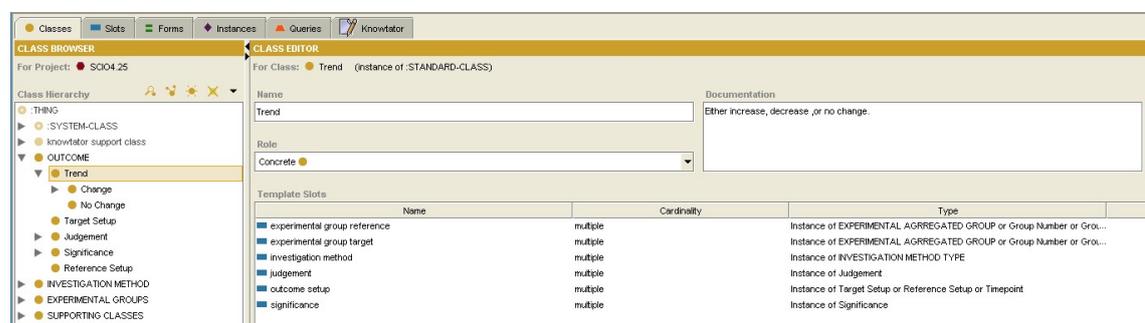


Abb. 19 **Die Attribute der Klasse Outcome.** Die markierte Unterklasse Trend ist die Schlüsselklasse und Mindestanforderung zur adäquaten Beschreibung eines Ergebnisses. Sie zeigt unter Anwendung einer bestimmten Untersuchungsmethode an, ob eine Therapie eine Veränderung induziert hat oder nicht. Die experimentelle Zielgruppe (*experimental group target*) und experimentelle Referenzgruppe (*experimental group reference*) sind Attribute der Hauptklasse *Outcome* und beinhalten das aggregierte Wissen des Material- und Methodenteils. Hinzu kommen die Attribute *significance*, *judgement* und *outcome setup*, die Kopien der gleichnamigen Unterklassen der Hauptklasse *Outcome* sind.

Zusammengefasst entspricht die Klasse *Outcome* der äußersten und größten Schale des zwiebelschalenartigen Aufbaus der horizontalen Wissensstruktur der SCIO. Die Klasse *Outcome* repräsentiert alle Konzepte einer präklinischen Studie und entspricht somit dem Ziel der Ontologie eine präklinische Studie in ihrer Gesamtheit übersichtlich und detailliert darzustellen. Die zu den *Outcomes* gehörenden Attribute können mit dem Konzept einer Datenkomprimierung verglichen werden. Abb. 19 zeigt, wie alle Attribute aus den Schlüsselklassen der einzelnen Hauptklassen im Sinne einer Verdichtung von Datenmengen den Attributen der Klasse *Outcome* zugeordnet werden. Bei diesem Prozess handelt es sich um eine verlustfreie Komprimierung, da alle

Originaldaten erhalten bleiben und ihr Ursprung auf die einzelnen Hauptklassen zurückverfolgt werden kann.

2.5 Implementierung und Evaluation der Ontologie

2.5.1 Anwendungsrichtlinien der Ontologie

Die SCIO basiert auf Konzepten zur Repräsentation der präklinischen Domäne über traumatischen Rückenmarksverletzungen. Im Wesentlichen besteht sie aus einem Gefüge aus Oberklassen, Unterklassen und Attributen, die den experimentellen Aufbau, die verwendeten Methoden und Ergebnisse einer Studie identifizieren und angemessen reflektieren. Die Kernaufgabe der SCIO ist die allumfassende, komplette Repräsentation einer präklinischen Studie, um infolgedessen einen Einblick auf die Signifikanz der präsentierten Ergebnisse zu gewährleisten. Hieraus soll in einem selektiven Verfahren die Fortführung erfolgsversprechender präklinischer Studien in klinische Studien erfolgen. Nur eine universelle und einheitliche Einsatzmöglichkeit der Ontologie kann diese Vorhaben nachhaltig unterstützen. Für eine fachgerechte und einheitliche Anwendung der SCIO wurden Annotationsrichtlinien erstellt. Mithilfe dieser Richtlinien soll der Ablauf des Annotationsprozesses von wissenschaftlichen Texten durch unterschiedliche Annotatoren erleichtert werden. Gleichzeitig dienen die Richtlinien als präzise Anleitung zur Anwendung der unterschiedlichen Klassen und Attribute, um ein gewisses Maß an Übereinstimmung bei den Annotationen durch verschiedene Annotatoren zu gewährleisten. Der einheitliche Gebrauch der SCIO ist für ihre Validierung als funktionale und letztendlich erfolgreich einsetzbare Ontologie von äußerster Relevanz. Die Annotationsrichtlinien bestehen aus 17 Seiten, 5 Abschnitten und heben die Annotation von Hauptkonzepten hervor, indem sie beschreiben, wie und wann die unterschiedlichen Klassen zur Annotation verwendet werden sollen. Der erste Abschnitt ist eine konzise Zusammenfassung über die Notwendigkeit und Ziele der SCIO. Der zweite Abschnitt beschreibt den strukturellen Aufbau der Hauptklassen *Experimental Groups*, *Investigation Method*, *Outcome* und *Supporting Classes* mit ihren dazugehörigen Unterklassen. Der dritte Abschnitt beschreibt die allgemeinen Prinzipien des Annotationsprozesses bei fachgerechter Anwendung der SCIO. Hier soll den Annotatoren ein Grundverständnis für den Annotationsprozess vermittelt werden. In den Abschnitten vier und fünf wird die Sequenz bei der Annotation des Material- und Methodenteils, sowie den Ergebnissen aus wissenschaftlichen Texten beschrieben.

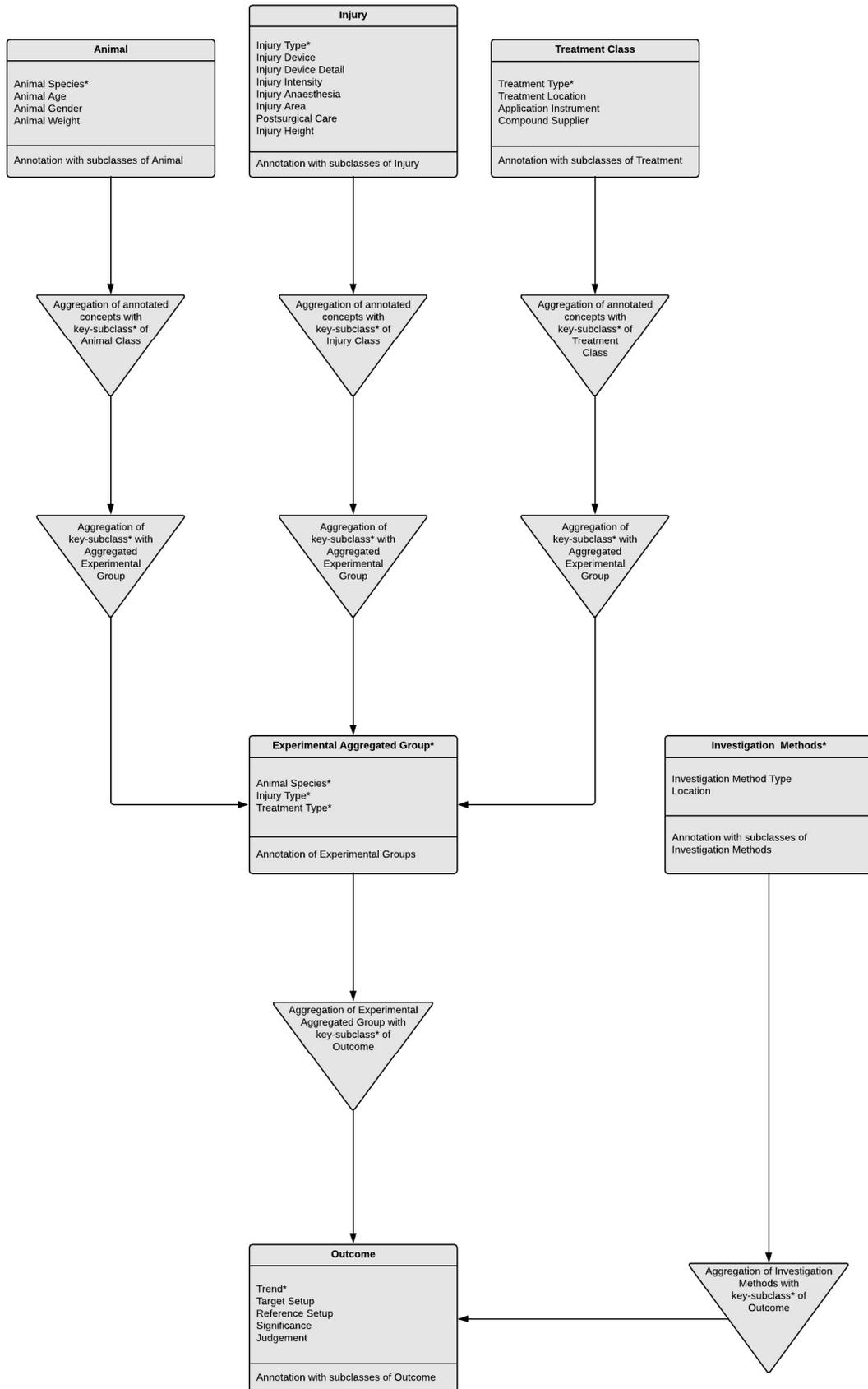


Abb. 20 **Das Annotationsschema mit SCIO.** Es zeigt die Reihenfolge der Annotation und Aggregation von extrahierten Konzepten aus einer wissenschaftlichen Publikation.

Wie aus Abb. 20 hervorgeht, beginnt der Annotationsprozess anlehnend am typischen Beschreibungsmuster des Material- und Methodenteils einer Publikation durch die manuelle Datenextraktion der experimentellen Gruppen. In diesem Zusammenhang erfolgt die Annotation des Tiermodells (*Animal*), gefolgt von der Annotation des Verletzungsmodells (*Injury*) und abschließend durch die Annotation des Therapiemodells (*Treatment*). Dabei werden unter Anwendung der Unterklassen die entsprechenden Textstellen einer Publikation markiert und somit annotiert. Die Annotation erfolgt auf höchster Detailstufe (Unterklassen), um für das Verständnis einer Studie und für die notwendigen Vergleiche zwischen den experimentellen Gruppen alle essentiellen Daten zu erfassen und eine adäquate Beurteilung der unterschiedlichen Behandlungskonzepte und den daraus hervorgehenden Resultaten zu ermöglichen. Bei den mit einem Asterisk gekennzeichneten Unterklassen in Abb. 20 handelt es sich um die Schlüsselklassen der unterschiedlichen Hauptklassen. Ihnen werden bei Anwendung der entsprechenden Unterklassen im Rahmen der Annotation die extrahierten Konzepte im Sinne einer Aggregation hinzugefügt und als Eigenschaften (slots) der Schlüsselklassen angezeigt. Dieses komprimierte Wissen wird gespeichert und kann auf Hauptklassen übertragen werden, die sich im Relationsgefüge der SCIO auf einer höheren Ebene befinden. Für die Annotation eines Konzeptes wird die passende Unterklasse ausgewählt und eine Textstelle mit ihr markiert. Alle Hauptklassen mit ihren dazugehörigen Unterklassen weisen zur Differenzierung eine individuelle Farbkodierung auf. Zur Veranschaulichung einer fachgerechten Anwendung der unterschiedlichen Klassen werden in den Annotationsrichtlinien Annotationsbeispiele gezeigt, um die zu extrahierenden Hauptkonzepte in den unterschiedlichen Textabschnitten hervorzuheben.

Aggregation of 'Animal':

Adult male CD1 mice weighing 25-30 g were housed separately and served animal chow and water.

Abb. 21 **Annotationsbeispiel mit der Hauptklasse *Animal***. Das den Annotationsrichtlinien entnommene Beispiel zeigt in gelber Farbkodierung die annotierten Konzepte der Hauptklasse *Animal* unter Anwendung der passenden Unterklassen. Die graue Farbkodierung entspricht der Hauptklasse *Supporting Classes* mit der die abstrakten Konzepte, wie Zahlenwerte, annotiert werden. In den Annotationsrichtlinien werden für alle Hauptklassen Annotationsbeispiele für eine optimale Datenextraktion aus wissenschaftlichen Texten gezeigt.

Das in Abb. 21 gezeigte Beispiel schildert die Annotation von vier Konzepten. Die Kernaussage des annotierten Textabschnitts ist, dass erwachsene, männliche CD1-Mäuse mit einem Körpergewicht von 25 bis 30 g für das beschriebene Experiment verwendet wurden. Diese Konzepte wurden mit den Unterklassen ihrer respektiven Oberklassen annotiert. Die Unterklasse *adult* als Teil der Hauptklasse *Animal Age*, die Unterklasse *male* als Teil der Hauptklasse *Animal Gender* und die Unterklasse *CD1 mice* der Hauptklasse *Animal Species* wurden für die Annotation dieses Textabschnittes verwendet. Innerhalb der Hauptklassen erfolgt die Aggregation aller annotierten Unterklassen mit den designierten Schlüsselklassen, die nach der Annotation den experimentellen Gruppen mittels Aggregation zugeordnet werden. Bei der Hauptklasse *Animal* ist die Schlüsselklasse die Unterklasse *Animal Species*, respektive die in Abb. 21 annotierte Spezies *CD1 mice*. Ihr werden im nächsten Schritt alle zur Hauptklasse *Animal* gehörenden und annotierten Konzepte durch Aggregation zugeordnet.

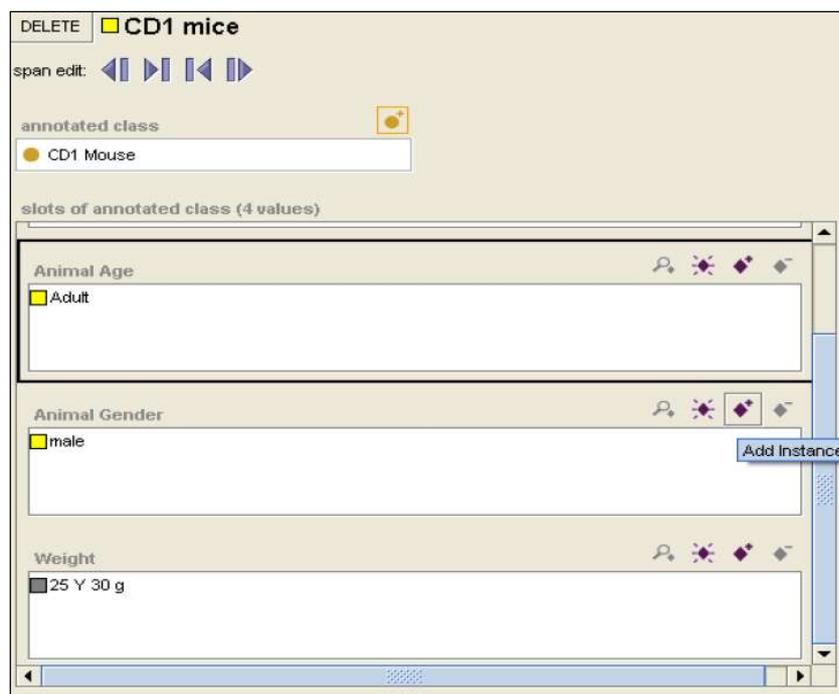


Abb. 22 **Aggregation der Klasse *Animal***. Die annotierte Textstelle *CD1 mice* (oben links mit gelber Farbkodierung) wird, wie in der unter ihr stehenden Rubrik *annotated class*, mit der zur Schlüsselklasse *Animal Species* gehörenden Unterklasse *CD1 Mouse* annotiert. Anschließend werden ihr alle annotierten Unterklassen, die in der Rubrik *slots of annotated class* als Attribute (slots) im Knowtator-Plugin des Annotationsprogramms Protegé zugeordnet. Diese annotierten Klassen sind nun Eigenschaften der Schlüsselklasse *Animal Species* und werden beim Aufrufen der annotierten Textstelle „*CD1 mice*“ im *Plug-in* als Zusammenfassung angezeigt.

In Abb. 22 werden in der Rubrik *slots of annotated classes* die vorgegebenen Attribute *Animal Age*, *Animal Gender* und *Weight* in grauer Überschrift angezeigt und können als eine in die Ontologie eingebaute Annotationsrichtlinie angesehen werden, da sie die Attribute vorgeben, die der Schlüsselklasse aus dem annotierten Text zuzuordnen sind. Hierbei werden diese nach vorausgegangener Annotation mit den entsprechenden Unterklassen der Hauptklasse *Animal* im Knowtator-Plugin in der Rubrik *slots of annotated class* als Eigenschaften der Schlüsselklasse *Animal Species* aggregiert und angezeigt. Dieses anhand der Hauptklasse *Animal* beschriebene Konzept der Annotation und Aggregation wird bei allen weiteren Hauptklassen appliziert, so dass im experimentellen Abschnitt einer wissenschaftlichen Publikation die Schlüsselklassen der Hauptklassen *Animal (Animal Species)*, *Injury (Injury Type)* und *Treatment (Compound)* den unterschiedlichen experimentellen Gruppen zugeordnet werden können. In den Annotationsrichtlinien werden nicht für alle existierenden Unterklassen Annotationsbeispiele erläutert, allerdings wird das generelle Vorgehen bei der Annotation anhand von unterschiedlichen Beispielen gezeigt, die auf die Anwendung anderer Klassen übertragen werden können. Abb. 23 zeigt das für jede Klasse der Ontologie angelegte Dokumentationsfeld, das eine konzise Beschreibung ihrer Funktion wiedergibt.

The screenshot shows a 'CLASS EDITOR' window. At the top, it says 'For Class: ● Animal Species (instance of :STANDARD-CLASS)'. Below this, there are two main sections: 'Name' and 'Documentation'. The 'Name' section has a text input field containing 'Animal Species'. The 'Role' section has a dropdown menu with 'Concrete ●' selected. The 'Documentation' section contains the text: 'Animal species, i.e. mouse, rat, monkey etc. (singular) Class of Aggregation within the animal paradigm.'

Abb. 23 **Beispiel der Dokumentation von Klassen.** In den Annotationsrichtlinien wird auf die in der Ontologie vorhandene Dokumentation der unterschiedlichen Klassen verwiesen. Hierbei handelt es sich um prägnante Beschreibungen der unterschiedlichen Klassen, die als Unterstützung bei der Anwendung von Klassen im Rahmen der Annotation dienen sollen.

In den Annotationsrichtlinien wird die Reihenfolge des Annotationsverfahrens für jede Hauptklasse im Einzelnen beschrieben. Hierbei wird die Markierung von Textstellen zur Annotation von Konzepten und die Aggregation von annotierten Unterklassen genau erklärt, um eine einfache Anwendung der Software zu gewährleisten. Nach der

Annotation der Klasse *Animal* wird die Klasse *Injury* annotiert und in den Richtlinien anhand von Textbeispielen und Abbildungsbeispielen wie in Abb. 24 erläutert.

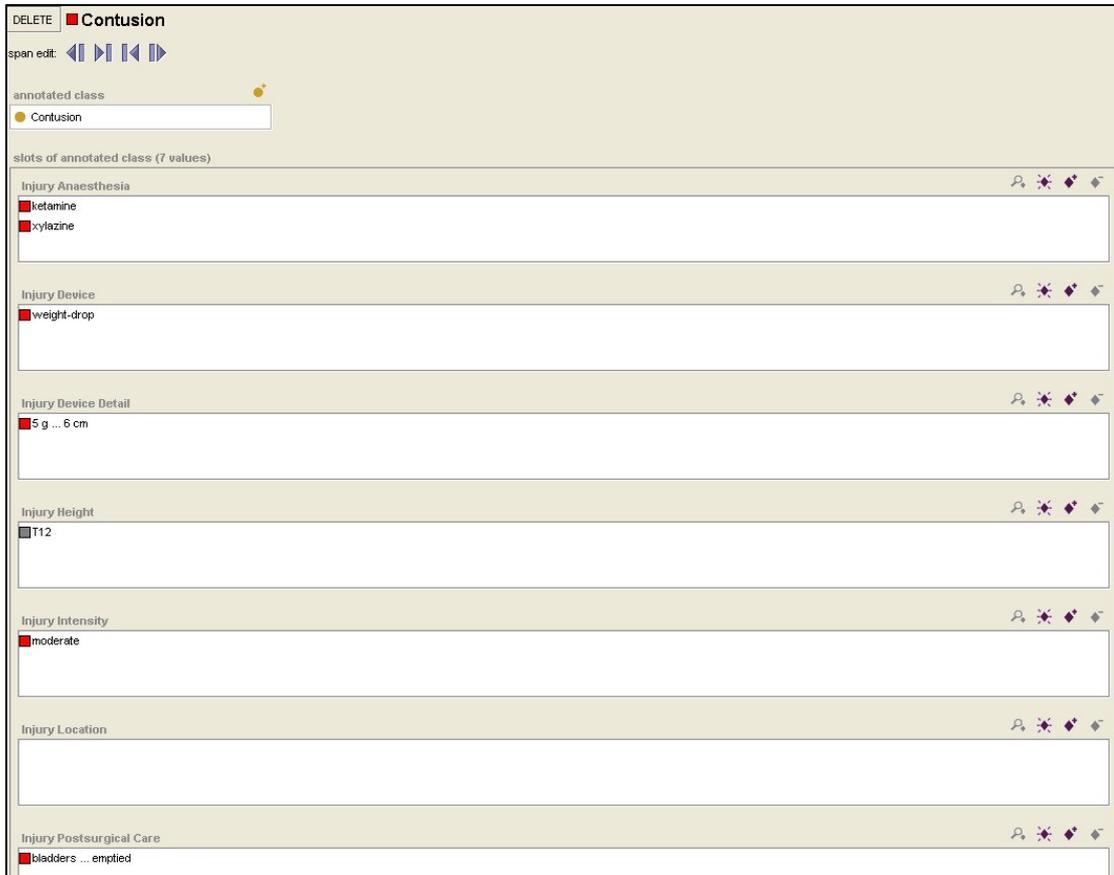


Abb. 24 **Aggregation der Klasse *Injury***. Mit der Unterklasse *Contusion* der Hauptklasse *Injury* Type wurde die experimentell verursachte Kontusion im Text annotiert. In der Rubrik *slots of annotated class* werden alle aggregierten Unterklassen der Hauptklasse *Injury* als Attribute (slots) angezeigt. Zusammenfassend wird beim Aufrufen der Verletzungsart im Text angezeigt welche Anästhetika präoperativ verabreicht, welche Apparatur für die experimentell verursachte Kontusion verwendet und welche Einstellungen an ihnen vorgenommen wurde. Der genaue anatomische Abschnitt des verletzten Rückenmarks und die Intensität der verursachten Kontusion werden in diesem Beispiel ebenfalls angezeigt. Die Unterklasse *Injury Location* wurde im Text nicht erwähnt und konnte somit nicht annotiert und als Eigenschaft von *Injury* Type aggregiert werden.

Wie bereits zuvor erwähnt, ist in Abb. 24 zu erkennen, dass die Hauptklasse *Injury* im Vergleich zur Hauptklasse *Animal* mehr Unterklassen mit sich führt. Dies ist darauf zurückzuführen, dass dieser Abschnitt des experimentellen Aufbaus tendenziell ausführlicher beschrieben wird als andere Abschnitte, was unter anderem an der Vielzahl der Läsionsmethoden liegt. Aus den Annotationsrichtlinien gehen die Hauptklassen *Animal* und *Injury* als Voraussetzungen des experimentellen Ablaufes

hervor. Im nächsten Abschnitt der Annotationsrichtlinien wird die Annotation der Hauptklasse *Treatment* beschrieben. Sie ist die letzte Komponente des experimentellen Ablaufes, die für eine präzise Beschreibung der unterschiedlichen experimentellen Tiergruppen von Bedeutung ist. Die nächste Abbildung zeigt die Aggregation der Klasse *Treatment*.

Abb. 25 **Aggregation der Klasse *Treatment***. Die Unterklasse *Compound* der Hauptklasse *Substance* wurde zur Annotation des verwendeten Wirkstoffes *Methylprednisolone* (oben links in grüner Farbkodierung) verwendet. Alle weiteren annotierten Konzepte wurden mit Unterklassen der *Supporting Classes* annotiert und der Schlüsselklasse *Compound* als Eigenschaften hinzugefügt.

Wie in Abb. 25 zu erkennen ist, konnten lediglich drei von sechs möglichen Attributen der Schlüsselklasse *Compound* hinzugefügt werden, was darauf zurückzuführen ist, dass die drei fehlenden im Text nicht erwähnt werden und somit nicht annotiert werden können. Der Wegfall von zu annotierenden Konzepten ist kein Versäumnis, was durch die Annotation von ähnlichen Konzepten ausgeglichen werden muss. Dies ist ein wichtiger Grundsatz, der aus den Annotationsrichtlinien hervorgeht und fehlerhafte Annotationen vermeiden soll. Nach erfolgter Annotation der Hauptklasse *Animal*, *Injury*

und *Treatment*, durch Anwendung der entsprechenden Unterklassen, beschreibt die Annotationsrichtlinie die korrekte Zuordnung der Schlüsselklassen zu den passenden experimentellen Gruppen. Hierbei werden alle Schlüsselklassen mitsamt ihren Eigenschaften einer experimentellen Gruppe zugeordnet. Zusammenfassend gehen somit aus der experimentellen Gruppe alle aggregierten Schlüsselklassen mit ihren Attributen hervor. Abb. 26 zeigt die Anwendung dieses Verfahrens.

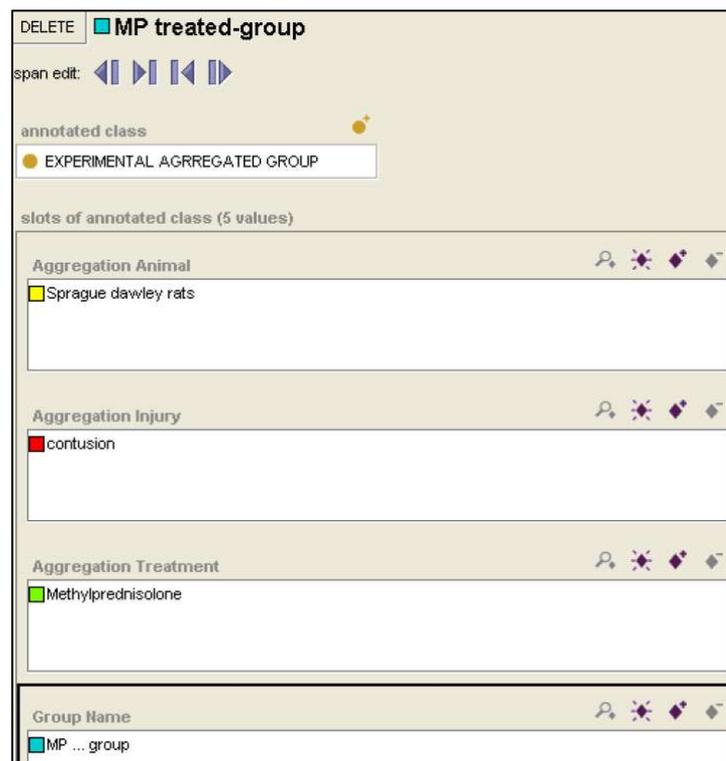


Abb. 26 **Aggregation der Klasse *Experimental Aggregated Group***. Aus der annotierten experimentellen Gruppe *MP treated-group* (in türkiser Farbkodierung oben links) geht nun hervor, dass es sich bei den Versuchstieren um Sprague Dawley rats handelt, denen nach einer experimentell verursachten Kontusionsverletzung des Rückenmarks der Wirkstoff Methylprednisolone als Therapie verabreicht wurde. Zusätzlich wird in der Rubrik *Group Name* angezeigt, dass die annotierte experimentelle Gruppe *MP-treated group* in anderen Textabschnitten als *MP group* bezeichnet wird. Alle hier aggregierten Schlüsselklassen führen ihre Eigenschaften mit sich. Die Klasse *Experimental Aggregated Group* wird wiederum zu einer Eigenschaft eines Ergebnisses, wenn unterschiedliche experimentelle Gruppen miteinander verglichen werden.

Nach Konsolidierung der annotierten Konzepte des Material- und Methodenteils erfolgt die Annotation der Ergebnisse. In diesem Abschnitt werden die unterschiedlichen experimentellen Gruppen und die mit ihnen verbundenen Therapien im Rahmen eines

direkten Vergleiches miteinander in Zusammenhang gesetzt. Die Vorgehensweise beim Vergleich der experimentellen Gruppen basierend auf einer bestimmten Untersuchungsmethode wird in den Annotationsrichtlinien beschrieben und anhand von Screenshots aus der Protegéé-Software erläutert. Das annotierte Ergebnis wird, wie zuvor erwähnt, durch die Annotation einer Zunahme (*increase*), Abnahme (*decrease*) oder einer ausbleibenden Veränderung (*no change*) annotiert. Diese Aussagen beziehen sich auf einen Richtwert, der durch die entsprechende Untersuchungsmethode charakterisiert wird. Ein Ergebnis wird gemäß der Annotationsrichtlinien zunächst durch die Annotation der Kernaussage beschrieben. Dieser Kernaussage werden die Untersuchungsmethode, die unterschiedlichen experimentellen Gruppen und die Signifikanz des festgestellten Resultates zugeordnet.

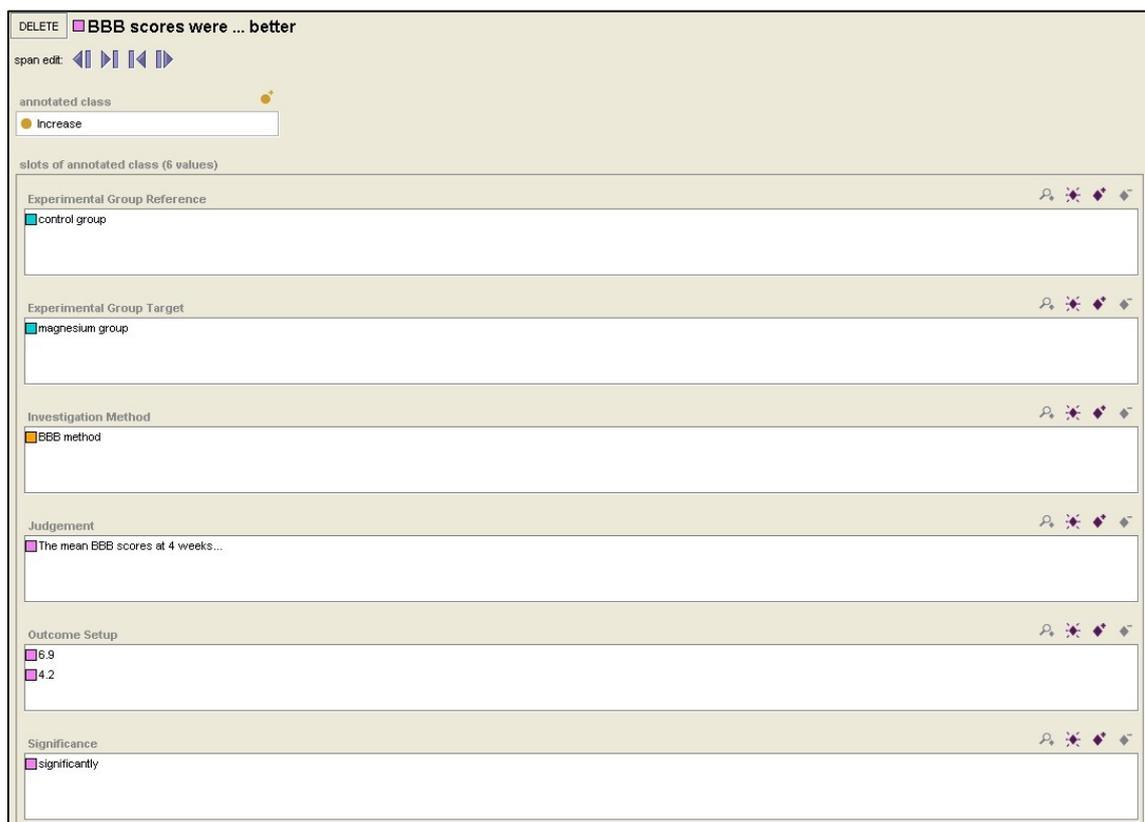


Abb. 27 **Aggregation der Klasse Outcome.** Dieser Screenshot aus der Protegéé-Software zeigt das annotierte Ergebnis „BBB scores were...better“ (oben links), das mit der Unterklasse *Increase* der Hauptklasse *Outcome* annotiert wurde. Der *BBB score* entspricht einer funktionalen Untersuchung der motorischen Kapazitäten von Versuchstieren. Wie aus der Rubrik *slots of annotated classes* hervorgeht, wurden die motorischen Eigenschaften der Versuchstiere zwischen einer mit Magnesium behandelten Zielgruppe und einer unbehandelten

Kontrollgruppe als Referenz vorgenommen. Hierbei erlangte die Zielgruppe einen niedrigeren, respektive einen besseren Punktestand als die Kontrollgruppe, wie in der Rubrik *Outcome Setup* angezeigt wird. Das Ergebnis wird als signifikant eingestuft, somit muss ein P-Wert von kleiner als 0,05 vorliegen. Es handelt sich um ein aussagekräftiges Ergebnis.

Die in Abb. 27 miteinander verglichenen experimentellen Gruppen beinhalten die aus dem Material- und Methodenteil extrahierten Informationen. Somit geht aus einem Ergebnis hervor, durch welche Tierspezies eine experimentelle Gruppe vertreten wird, wie sie verletzt und behandelt wurde. Hieraus können Schlüsse auf den Erfolg einer Therapie gezogen werden.

Die Annotationsrichtlinie beschreibt die drei Stufen der Datenkomprimierung, oder Wissensaggregation, um das in seiner großen Quantität vorhandene Wissen auffindbar und auf weitere Oberklassen übertragbar zu machen. In der ersten Instanz erfolgt die Annotation und Aggregation auf der Ebene der Hauptklassen Animal, Injury und Treatment. Die hieraus hervorgehenden Schlüsselklassen werden mit ihren Eigenschaften auf dem Niveau der experimentellen Gruppen in zweiter Instanz aggregiert, um abschließend in dritter Instanz beim Vergleich der unterschiedlichen experimentellen Gruppen basierend auf einer Untersuchungsmethode mit einem Ergebnis aggregiert zu werden. So werden die zuvor akkumulierten Daten in komprimierter Form für eine korrekte und zusammenhängende Präsentation der Ergebnisse bereitgestellt. Letztendlich führt die Annotation und Aggregation von Klassen zu einer adäquaten Beschreibung aller Komponenten einer präklinischen Studie, die durch die SCIO präsentiert werden. Gleichzeitig wird ein Verlust von kritischer Information für das Verständnis einer präklinischen Studie vermieden.

Durch die Einhaltung einer Sequenz soll eine einheitliche Anwendung der Ontologie durch verschiedene Anwender unterstützt werden. Darüber hinaus soll den Annotatoren vermittelt werden, nach welchen Konzepten sie im Text Ausschau halten müssen, um die Wissensessenz einer Studie einzufangen. Die Annotationsrichtlinien entsprechen einem Verbindungsglied zwischen der Ontologie und den Annotatoren, das den Ablauf der manuellen Datenextraktion erleichtert. Sie beschreibt in getrennten Abschnitten die Annotation der unterschiedlichen Hauptklassen und Unterklassen anhand von Beispielen, indem Screenshots bei der Anwendung der SCIO mittels der Protégée- und Knowtator-Software präsentiert werden. Das Hauptziel der Annotationsrichtlinien ist eine objektive Annotation und Datenextraktion aus wissenschaftlichen Texten.

Zusätzlich bilden sie zusammen mit der Ontologie einen Vorläufer für die automatische Datenextraktion. Zur Beurteilung des Wirkgrades der SCIO musste sie in Kombination mit den Richtlinien von unterschiedlichen Annotatoren gleichzeitig angewendet werden.

2.3.2 Das Annotationsprojekt

Zur Validierung der SCIO wurde ein Annotationsprojekt mit zwei Annotatoren durchgeführt. Bei der Auswahl der Annotatoren wurde berücksichtigt, dass es sich um Individuen handelte, die mit der Materie der traumatischen Rückenmarksverletzungen in der präklinischen Umgebung vertraut waren. Zwei Diplom-Biologen, die für das Labor für Molekulare Neurobiologie an der Neurologischen Klinik des Universitätsklinikums der Heinrich-Heine-Universität arbeiteten und zuvor an der Ausführung von experimentellen Studien dieser Art in der Tierversuchsanstalt der Heinrich-Heine-Universität beteiligt waren, wurden für das Annotationsprojekt eingesetzt.

Den Annotatoren wurde die Protegée 3.3.1-Software mit dem Knowtator-Plugin auf zwei Notebooks mit dem Windows 8 Betriebssystem installiert. Die fertige und aktualisierte Version der Spinal Cord Injury Ontology wurde ebenfalls in der Protegée-Software auf zwei Notebooks installiert. In den Knowtator-Plugin wurden auf jedem Notebook jeweils drei Publikationen über TRV importiert und geöffnet. Die Herangehensweise beim Import der Publikationen in Form von TXT-Dateien in das Plugin wurde den Annotatoren demonstriert, da dies im weiteren Verlauf durch sie selbstständig erfolgen musste. Es folgte eine Präsentation der Annotationsrichtlinien mit einem parallelen Tutorial über die Anwendung des Knowtator-Plugins anhand von praktischen Beispielen. In diesem Zusammenhang wurden Beispielannotationen durchgeführt, um die Annotatoren mit der Handhabung der Software vertraut zu machen. Gleichzeitig erfolgte neben der Erklärung der Annotationsregeln eine Diskussion mit Fragen und Antworten, um eine möglichst effektive Anwendung der Ontologie zu unterstützen. Zunächst wurden lediglich kurze Textabschnitte des Material- und Methodenteils annotiert. Anschließend erfolgte eine Übung zum Verständnis der Aggregation von annotierten Konzepten, um die experimentellen Gruppen in ihrer Gesamtheit adäquat zu repräsentieren. Die gleichen Schritte wurden für die Annotation der Ergebnisse erläutert. Die erste von den drei wissenschaftlichen Publikationen, die in das Knowtator-Plugin importiert wurden, wurde im Rahmen eines

praktischen Tutorials gemeinsam annotiert. Die in den Annotationsrichtlinien beschriebene Sequenz beim Annotationsablauf und das Hauptaugenmerk der zu extrahierenden Konzepte wurde besprochen und anschließend diskutiert. Die letzten beiden wissenschaftlichen Texte wurden von den Annotatoren selbstständig und getrennt voneinander unter zur Hilfenahme der Annotationsrichtlinien annotiert. Im Anschluss wurden die Annotationen miteinander verglichen und erneut diskutiert. Abschließend wurden den Annotatoren die ersten fünf von insgesamt zweiundzwanzig, selbstständig zu annotierenden Publikationen in die Knowtator-Software importiert. Die Annotation dieser Publikationen sollte im weiteren Verlauf im privaten Umfeld der Annotatoren erfolgen und nach ihrer Fertigstellung an das CNR per E-Mail geschickt werden.

Während der Annotation der ersten fünf Publikationen durften die Annotatoren schriftliche Fragen bezüglich der Softwarehandhabung stellen, die zeitgerecht beantwortet wurden. Nachdem die Annotatoren die Annotationen der ersten fünf Publikationen absolvierten, wurden die nächsten fünf Publikationen im TXT-Format per E-Mail an sie verschickt. Dieser Prozess wiederholte sich bis alle zweiundzwanzig Publikationen annotiert wurden. Im Anschluss an das Projekt erfolgten Analysen zur Beurteilung der Übereinstimmung bei den Annotationen aller Publikation mit SCIO.

3 Ergebnisse

3.1 Die relationale Struktur und Anwendung der SCIO

Das zentrale Konzept der SCIO ist die Darstellung eines Ergebnisses als eine zusammenhängende, relationale Struktur bestehend aus verschiedenen Klassen und Hierarchien, um Rückschlüsse auf die verwendeten Materialien, die Methoden und die Signifikanz eines Ergebnisses zu ziehen. Aufgrund der unterschiedlichen, auf sich aufbauenden Aggregate von Wissenseinheiten, entsteht ein konfluierender Wissensstrom, der ein Ergebnis mit all seinen Komponenten, respektive Oberklassen, Unterklassen und Eigenschaften im Ganzen darstellt, ohne dass relevante Informationen über den experimentellen Aufbau und Ablauf verloren gehen. So kann das komplexe Wissen, dass aus einer präklinischen Tierstudie über traumatische Rückenmarksverletzungen hervorgeht, übersichtlich und gleichzeitig auf einer hohen

Detailstufe abgebildet werden, um es für seine Erweiterbarkeit in einer klinischen Studie beurteilen zu können. Abb. 28 illustriert das Informationsnetzwerk der SCIO.

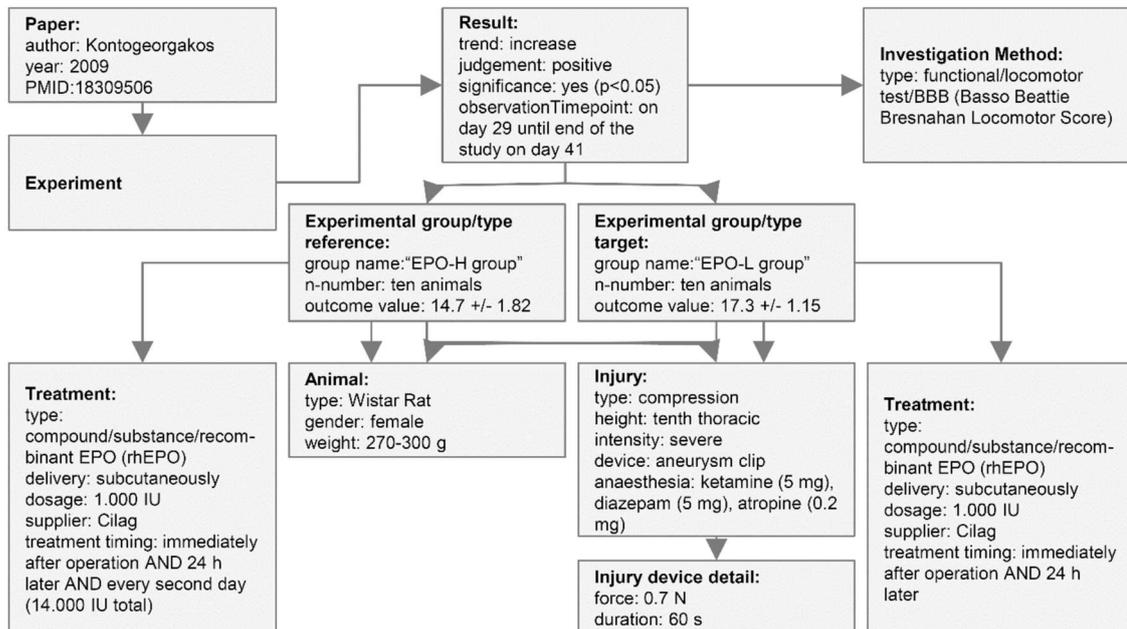


Abb. 28 **Beispiel der Extraktion von Informationen mit SCIO.** Die relevante Information eines Experimentes aus einer wissenschaftlichen Publikation (Pubmed ID18309506) wird mit SCIO als Datennetzwerk abgebildet. Aus der dargestellten Publikation gehen eine Vielzahl an Ergebnissen hervor. Für eine bessere Übersicht wurde lediglich ein signifikantes Ergebnis in dieser Abbildung dargestellt.

Das zentrale Konzept der SCIO ist die Unterklasse *trend* der Hauptklasse *Result*, die das qualitative Ergebnis einer präklinischen Studie beschreibt. In Abhängigkeit von den zu beobachtenden Parametern einer Laboruntersuchung kann die Klasse *trend* eine Zunahme (*increase*), Abnahme (*decrease*) oder keine Veränderung (*no change*) anzeigen. In Abb. 28 wird im Rahmen einer funktionalen Untersuchungsmethode bei der Zielgruppe eine Verbesserung (*increase*) der motorischen Fähigkeiten festgestellt, die im Vergleich zur Referenzgruppe mit einer niedrigeren Dosis rhEPO behandelt wurde. Das Ergebnis (*Result*) zeigt, dass die Zunahme der wiedergewonnenen, motorischen Fähigkeiten signifikant ist. Die für dieses Experiment verwendeten Versuchstiere, angewendeten Verletzungsmethoden und Behandlungsmethoden gehen aus den zwei, miteinander verglichenen, experimentellen Gruppen hervor. Zusammenfassend kann konstatiert werden, dass die aus einem Ergebnis hervorgehenden Informationen über die Tierspezies, die experimentell induzierten Verletzungen, die Behandlungsmethode und über weitere Aspekte des experimentellen

Setups auf drei verschiedenen Niveaus der Klassenhierarchie aggregiert werden, um letztendlich einem Vergleich zwischen verschiedenen experimentellen Gruppen, und somit einem Ergebnis, zugeordnet zu werden. Dies gilt auch für extrahierte Informationen mit der Klasse *Investigation Methods*, die in wissenschaftlichen Texten oft auf die namentliche Erwähnung der angewendeten Untersuchungsmethode reduziert wird und keine weiteren detaillierten Auskünfte, außer die Lokalisation des untersuchten Rückenmarkabschnitts, über den methodologischen Ablauf einer Untersuchung vermitteln. Die Zeitpunkte bei den Untersuchungsmethoden sind nicht Teil der Klasse *Investigation Methods*, sondern Teil der Ergebnisse, da sie erst in den Ergebnissen von wissenschaftlichen Texten erwähnt werden. Die Klasse *Judgement*, respektive die Beurteilung eines Ergebnisses als „*positive*“, „*negative*“ oder „*no effect*“, basiert ausschließlich auf der persönlichen Meinung der Autorin oder des Autors einer wissenschaftlichen Publikation und gehört ebenfalls zu den Ergebnissen.

3.2 Analyse und Evaluation des Annotationsprojektes

Die zweiundzwanzig annotierten Publikationen wurden miteinander verglichen, indem die einzelnen Annotationen aus jeder Publikation in einer Excel-Tabelle eingepflegt und übersichtlich dargestellt wurden. Hierbei wurden neben den verwendeten Hauptklassen und Unterklassen die annotierten Textstellen angezeigt. Alle Annotationen mit den fünf Hauptklassen *Animal*, *Injury*, *Treatment*, *Experimental Groups* und *Outcome* wurden zunächst unabhängig voneinander verglichen. Abschließend erfolgte ein kumulativer Vergleich zur Feststellung der Übereinstimmung bei den Annotationen aller Publikationen. Diese Herangehensweise sollte mögliche Diskrepanzen bei den Annotationen durch unterschiedliche Personen offenlegen. Darüber hinaus sollte das Verhaltensmuster der Annotatoren bei der Datenextraktion evaluiert werden, um Rückschlüsse auf mögliche individuelle Fehlerquellen bei der Datenextraktion festzustellen. Außerdem sollten Rückschlüsse auf die Qualität der schriftlichen Darstellung des experimentellen Setups gezogen werden, die ggf. für eine fehlende Übereinstimmung bei der Annotation von wissenschaftlichen Veröffentlichungen verantwortlich sein könnten.

3.2.1 Der Inter-rater reliability-Wert (P_0)

Zur Beurteilung der beobachteten Übereinstimmung von Annotationen zwischen zwei Annotatoren kann der *Inter-rater reliability*-Wert (IRR-Wert) bestimmt werden. Er entspricht der prozentualen Übereinstimmung (P_0) von Annotationen und wird unter Anwendung der allgemeinen Prozentformel berechnet [22]:

$$P_0 = \frac{\text{Anzahl der übereinstimmenden Annotationen}}{\text{Gesamtzahl der Annotationen}} \times 100$$

Nachdem die IRR-Werte für die annotierten Hauptklassen jeder Publikation einzeln bestimmt wurden, erfolgte eine kumulative Berechnung der IRR-Werte für jede Hauptklasse aus allen 22 annotierten Publikationen.

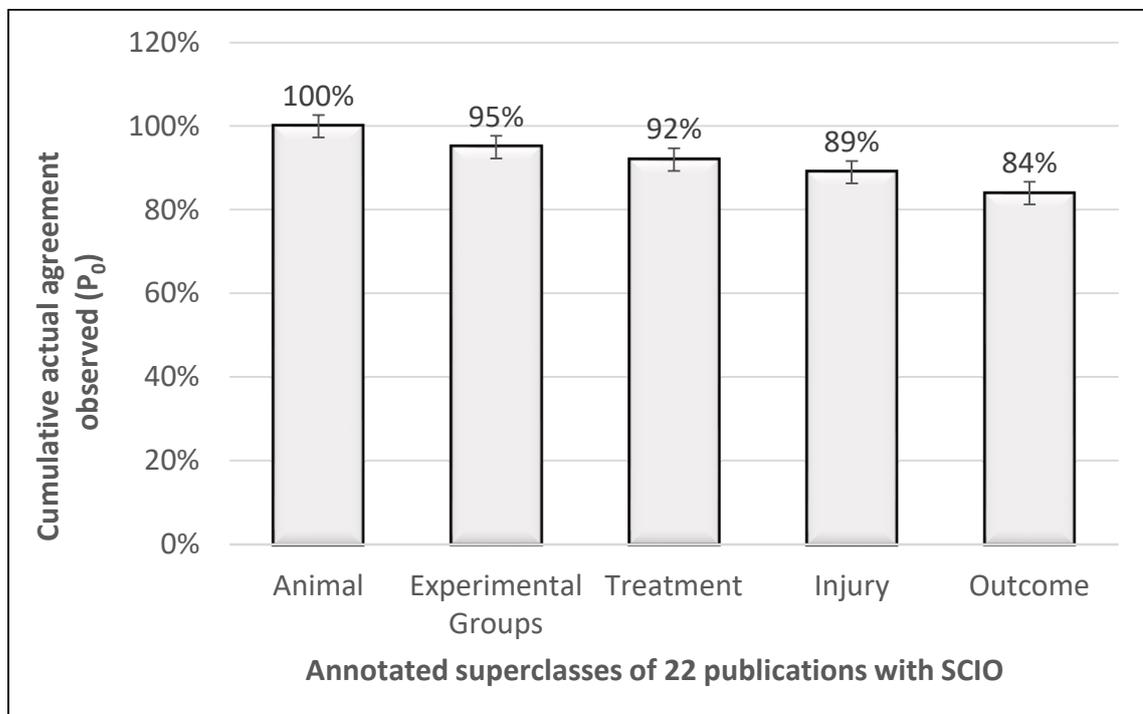


Abb. 29 **Kumulativer IRR-Wert (P_0) der Hauptklassen bei der Annotation mit SCIO.** Eine Zusammenfassung der prozentualen Übereinstimmung bei den annotierten Hauptklassen aus 22 Publikationen mit SCIO durch zwei Annotatoren.

Aus Abb. 29 geht hervor, dass bei 22 annotierten wissenschaftlichen Publikationen mit SCIO eine perfekte Übereinstimmung bei der Annotation der Klasse *Animal* vorlag. Die Klassen *Experimental Groups*, *Treatment*, *Injury* und *Outcome* zeigten einen IRR-Wert von 95%, 92 %, 89% und 84 %, was einer substanziellen Übereinstimmung entspricht. Mithilfe der SCIO konnte somit ein hohes Maß an prozentualer Übereinstimmung bei der Datenextraktion erreicht werden.

3.2.2 Der Cohen's Kappa Koeffizient (κ)

Um die Genauigkeit und Beständigkeit von Annotationen zwischen zwei Annotatoren festzustellen, muss in Erwägung gezogen werden, dass Übereinstimmungen zufällig entstehen können und gleichzeitig sichergestellt werden muss, dass die von den Annotatoren extrahierten Daten eine korrekte Repräsentation der gemessenen Variablen sind [22]. Der Cohen's Kappa Koeffizient (κ) gewährleistet genau dies, indem er die Konstanz der Annotationen von zwei unterschiedlichen Annotatoren überprüft und zufällige Übereinstimmungen, aufgrund von individuellen Unsicherheiten oder Fehlentscheidungen, statistisch miteinbezieht [32]. Die Proportion zufälliger Übereinstimmungen (P_e) errechnet sich aus den dichotomen Möglichkeiten, die entstehen, wenn beide Beurteiler ein Konzept annotieren (a), wenn beide Beurteiler ein Konzept nicht annotieren (d), wenn Beurteiler 1 ein Konzept annotiert, aber Beurteiler 2 nicht (c) und umgekehrt (b) [6]:

$$P_e = \frac{[(a + c) \times (a + b) + (c + d) \times (b + d)]}{(a + b + c + d)^2}$$

Die Wahrscheinlichkeit, dass beide Annotatoren ein Konzept mit der richtigen Klasse annotieren oder nicht annotieren, wird mit κ antizipiert und bestimmt. Somit können neben der allgemeinen Übereinstimmung mögliche, auf den Zufall zurückzuführende Fehlerquellen für eine allumfassende Auswertung in Betracht gezogen werden.

$$Kappa (\kappa) = \frac{P_0 - P_e}{1 - P_e}$$

Mit der Bestimmung des Cohen's Kappa Koeffizienten sollte beurteilt werden, ob die SCIO in Kombination mit den Annotationsrichtlinien eine adäquate Annotation mit wenigen Diskrepanzen zwischen den beiden Beurteilern ermöglichte.

<i>Kappa Statistic</i>	<i>Strength of Agreement</i>
<0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost Perfect

Abb. 30: **Das arbiträre Bewertungsschema des Cohen's Kappa Koeffizienten nach Landis, Koch 1977** – The M.jpg [17] Es dient als Richtwert bei der Einstufung des Koeffizienten.

Der Cohen's Kappa Koeffizient wurde für jede der zweiundzwanzig Publikation einzeln bestimmt und fasste die Übereinstimmung aller Annotationen einer Publikation zusammen. Alle für die Annotation einer Publikation verwendeten Klassen wurden für die Berechnung des Cohen's Kappa Koeffizienten verwendet, um das vermittelte Wissen der unterschiedlichen Autoren vollständig zu bewerten und miteinander zu vergleichen.

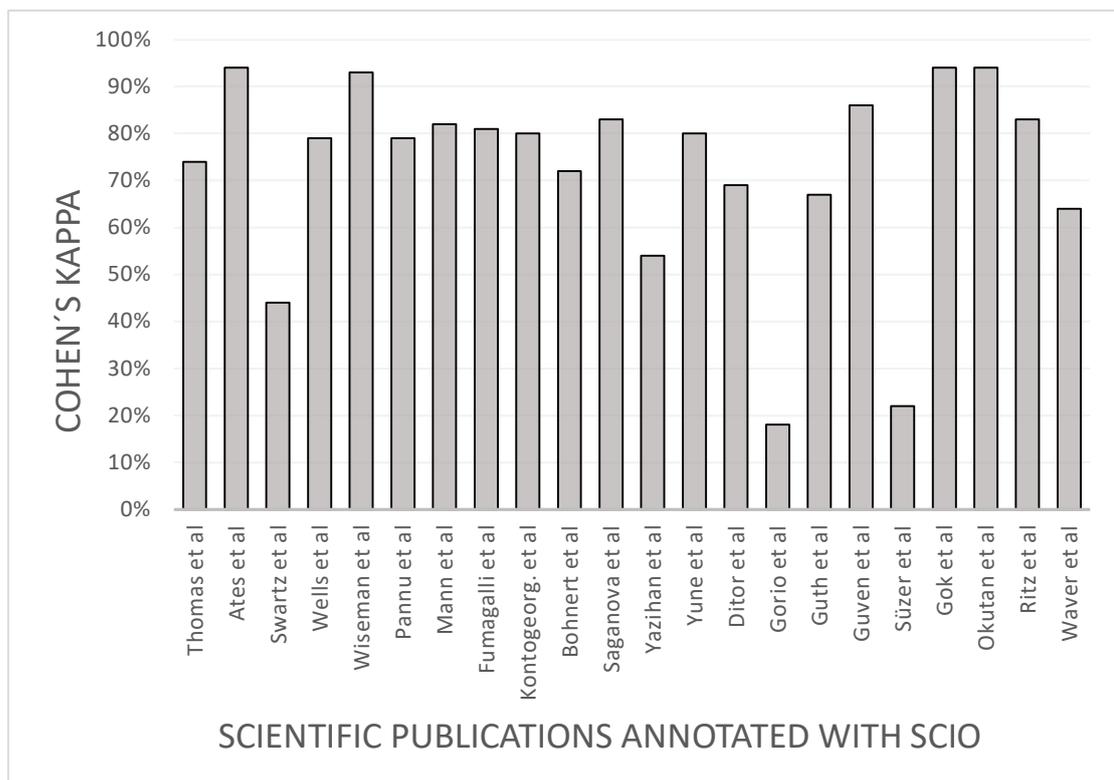


Abb. 31 **Cohen's Kappa Koeffizient der einzelnen Publikationen.** Bei neun Publikationen entsprach κ einer fast perfekten Übereinstimmung. Acht der zweiundzwanzig Publikationen waren mit einer substantziellen Übereinstimmung gleichzusetzen. Zwei Publikation wiesen eine moderate, respektive eine angemessene Übereinstimmung auf, während zwei weitere Publikation jeweils eine leichte und schlechte Übereinstimmung aufwiesen.

Wie aus Abb. 31 hervorgeht konnte, der Inhalt einer Publikation unter Anwendung der SCIO adäquat extrahiert und widergespiegelt werden. In den meisten Fällen konnten überdurchschnittlich hohe Übereinstimmungen bei den Annotationen erzielt werden, was auf die einfache und uniforme Anwendung der Ontologie in Kombination mit den Annotationsrichtlinien zurückzuführen ist. In Abb. 32 werden die unterschiedlichen Ausprägungen der erhobenen Übereinstimmungswerte miteinander verglichen.

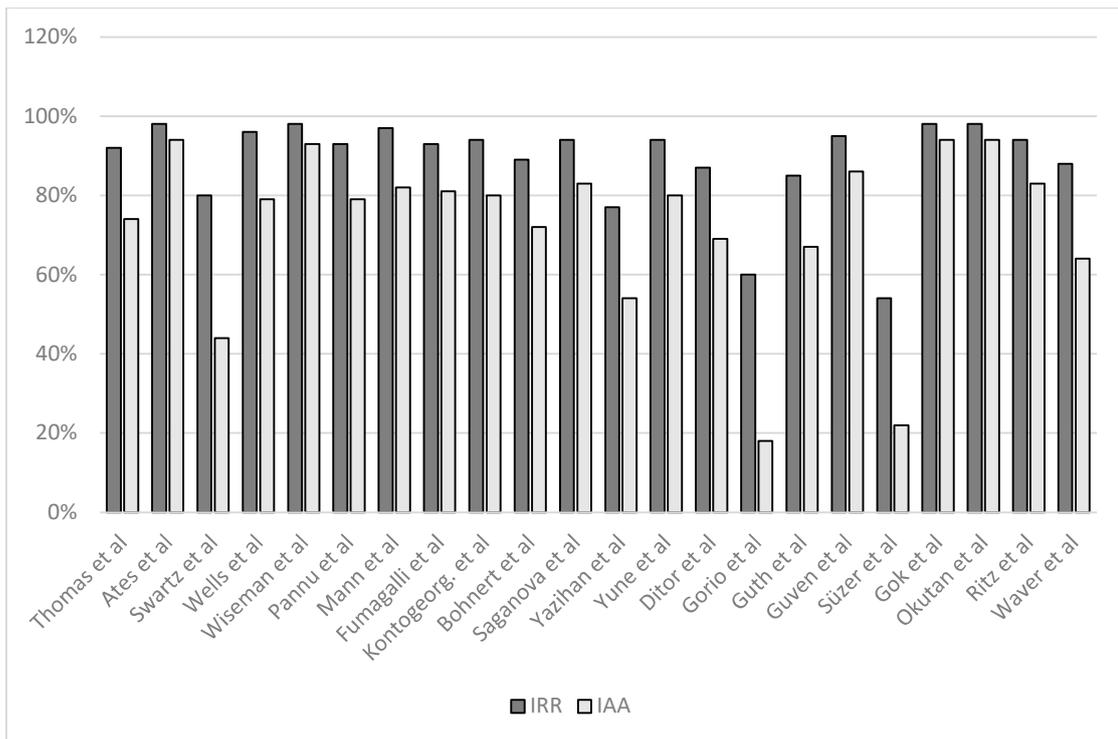


Abb. 32 Vergleich zwischen IRR-Wert und Cohen's Kappa Koeffizient (IAA). Der IRR-Wert zeigt im Vergleich zum Cohen's Kappa Koeffizienten, für dieselben 22 Publikationen, höhere Übereinstimmungswerte, da zufällige Übereinstimmungen nicht berücksichtigt werden.

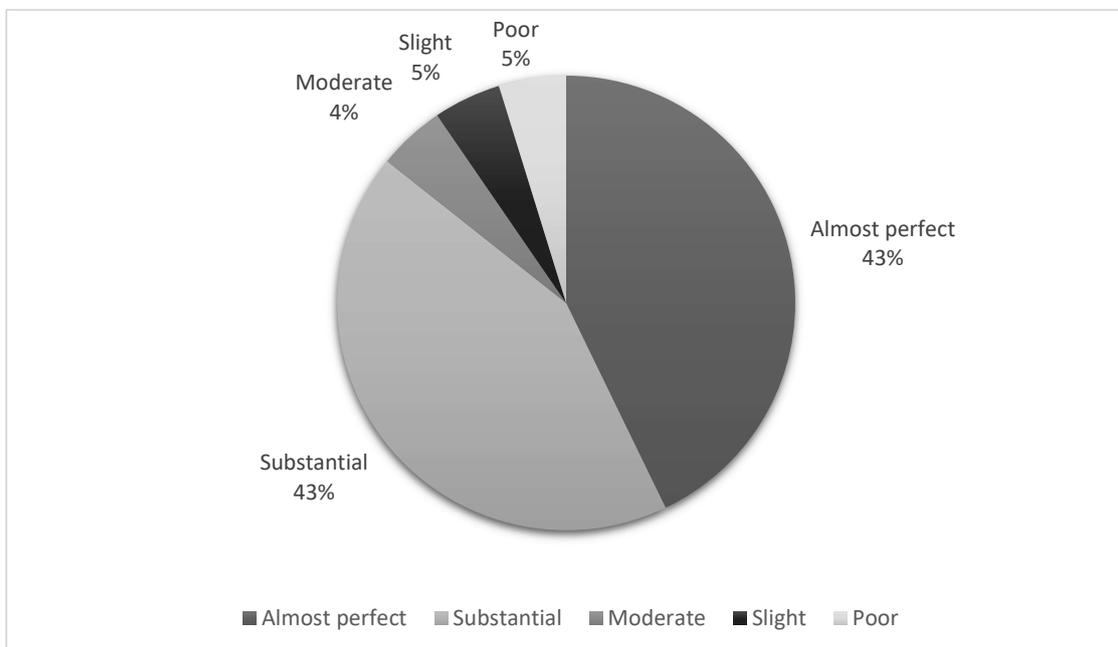


Abb. 33 Die Cohen's Kappa Koeffizienten der annotierten Publikationen. Die Datenextraktion mit SCIO zeigt, dass die Summe der fast perfekten und substanziellen Kappa Koeffizienten eine überdurchschnittliche Übereinstimmung von insgesamt 86 % ergeben.

Aus Abb. 33 geht hervor, dass durch die Applikation der SCIO auf wissenschaftliche Texte über präklinische Experimente zum Thema traumatische Rückenmarksverletzungen, eine hohe Übereinstimmung bei der Datenextraktion von signifikantem Ausmaß erreicht werden konnte. Somit konnte das Proof of Concept der Ontologie bewiesen werden. Alle Konzepte, die für das Verständnis und die Bewertung von präklinischen Studien notwendig sind, können im Rahmen der semiautomatischen Datenextraktion mit SCIO identifiziert und übersichtlich präsentiert werden. Sowohl die informellen als auch die formellen Kompetenzfragen können durch die Datenextraktion mit SCIO beantwortet werden. Somit ist die SCIO in der Lage eine präklinische Studie über TRV übersichtlich und vollständig darzustellen.

4 Diskussion

4.1 Erkenntnisse aus dem Annotationsprojekt

In Anbetracht der individuellen Interpretationsmöglichkeiten von schriftlich beschriebenen Zusammenhängen unterschiedlicher Konzepte, ist die alleinige Anwendung der Ontologie bei der Datenextraktion eine Herausforderung. Die Implementierung der Annotationsrichtlinien in Kombination mit der SCIO verleiht den Annotatoren einen Fokus und ein Gespür für ihre korrekte und zielgerichtete Anwendung bei der Datenextraktion aus wissenschaftlichen Publikationen. Das Hauptziel der Annotationsrichtlinien, eine uniforme Anwendung der Ontologie zu unterstützen, um möglichst hohe Übereinstimmungen bei den Annotationen zwischen zwei Anwendern zu erzeugen, wurde durch die hohen Cohen's Kappa Koeffizienten beim Annotationsprojekt bestätigt. Dennoch konnten Diskrepanzen bei den Annotationen festgestellt werden. Die in Abb. 29 dargestellte, kumulative, prozentuale Übereinstimmung bei der Annotation mit den unterschiedlichen Hauptklassen und den dazugehörigen Unterklassen zeigt, dass einige Klassen größere Übereinstimmungen aufwiesen als andere. Dies ist auf die Größe der Klassen und deren Darstellungsumfang in Publikationen zurückzuführen, was zum einen zu uneindeutigen Darstellungen einer Studie durch Autoren oder zum anderen zu fehlenden Übereinstimmungen bei der Annotation von Texten führen kann. Die Annotationen mit der Klasse *Animal* ergaben eine prozentuale Übereinstimmung von 100%, was auf ihre überschaubare Größe und ihrer konzisen und klaren Beschreibung im Material- und Methodenteil von

Publikationen zurückzuführen ist. Demnach erleichtert eine übersichtliche Darstellung von Konzepten ihre korrekte und vollständige Datenextraktion mit hoher Übereinstimmung beim Annotationsverfahren. Im Umkehrschluss kann die fehlende Übereinstimmung von beurteilenden Personen auf ein schwaches und zugleich unzureichend beschriebenes Studiendesign zurückzuführen sein [22]. Im Vergleich zur Klasse *Animal* sind die reduzierten Übereinstimmungen der anderen Klassen, auf ihre größere Komplexität zurückzuführen. Dies erfordert eine eindeutige und genaue Beschreibung des Sachverhaltes durch die Autoren, um komplexe Zusammenhänge unmissverständlich zu schildern. Annotatoren, sowie die Leser von Publikationen müssen in der Lage sein, Zusammenhänge zu verstehen, um essenzielle und für das Verständnis notwendige Daten mithilfe der Ontologie extrahieren zu können. Abb. 29 zeigt, dass die Klasse *Outcome* mit einem IRR-Wert von 84% das geringste Ausmaß an prozentualer Übereinstimmung zwischen den Annotatoren aufwies. Auch wenn dieser Wert einer überdurchschnittlichen Übereinstimmung entspricht, so zeigen sich in diesem Abschnitt die größten Unregelmäßigkeiten im Annotationsprojekt. Bei der Auswertung des Annotationsprojektes ging aus den Protokollen hervor, dass insbesondere bei der Darstellung der Ergebnisse eine Vielzahl an Unregelmäßigkeiten auftraten. Vergleiche zwischen Zielgruppen und Referenzgruppen wurden von Autoren nicht eindeutig beschrieben, darüber hinaus wurden als signifikant angegebene Ergebnisse ohne einen begleitenden p-Wert zur Validierung präsentiert. Die Autoren einer bestimmten Publikation verwendeten signifikante Ergebnisse einer fremden Publikation, um der Präsentation der eigenen Ergebnisse einen größeren Ausdruck zu verleihen, was bei der Annotation zu Missverständnissen und letztendlich zu Diskrepanzen führte. Zusammenfassend konnte im Rahmen des Annotationsprojektes festgestellt werden, dass die Qualität der Annotationen neben individuellen Fehlern von Seiten der Annotatoren sehr stark mit der Qualität und der Präsentation des Inhaltes einer Publikation korrelierte. Im Allgemeinen wurde in den Publikationen der experimentelle Aufbau genauer beschrieben als die Ergebnisse, was sicherlich auf die notwendige Konsolidierung von verschiedenen Datenpunkten zur vollständigen Darstellung eines Ergebnisses zurückzuführen ist. Daraus resultierten lückenhafte und in Anteilen uneindeutige Beschreibungen durch Autoren, die den Annotationsprozess erschwerten. Dies erklärt die unterschiedlich hohen prozentualen Übereinstimmungen der verschiedenen Klassen und Cohen's Kappa Koeffizienten der annotierten Publikationen wie in den Abb. 31, 32 und 33 beschrieben. Zusammenfassend bilden die

SCIO, die Annotationsrichtlinien und die Annotatoren eine beeinflussbare Einheit, die auf den Inhalt einer wissenschaftlichen Publikation angewendet werden, um das externe Wissen systematisch einzufangen. Die Tatsache, dass die Ontologie in ihrem Aufbau der Struktur einer wissenschaftlichen Publikation entspricht, fördert eine gezielte Extraktion und Darstellung von Konzepten, die für einen verständlichen Überblick und eine genaue Beurteilung von Publikationen von großer Bedeutung sind. Die Konzepte der Ontologie können sich bei der Entwicklung von präklinischen Studien als hilfreich erweisen. Ihre Konzepte suggerieren bereits Schlüsselemente, die für die Aussagekraft und Reproduzierbarkeit einer Studie von großer Bedeutung sein können. In dieser Funktion dient die Ontologie z.Z. lediglich als konkomitante Unterstützung beim Studiendesign und nicht als Richtlinie. Aktuell gibt es keine Richtlinien für eine adäquate schriftliche Berichterstattung über den experimentellen Aufbau und über die Ergebnisse einer Studie [19]. In diesem Zusammenhang bietet es sich an festzulegen, welche Konzepte sowohl im Studiendesign als auch bei der Verfassung von Publikationen inkorporiert und präsentiert werden sollten, um die Signifikanz einer Studie nachhaltig beurteilen zu können. Hieraus entwickeln sich die für einen Evidenzgrad wichtigen Erfahrungswerte.

4.2 Erweiterbare Anwendungsbereiche der SCIO

Durch die Zurverfügungstellung einer Konzeptualisierung der SCI-Domäne in Form von relevanten Konzepten und Relationen, die in aktuellen experimentellen Studien von Bedeutung sind, legt die SCIO das Fundament für eine umfassende Wissensdatenbank, die den gesamten Korpus des erworbenen Wissens, das in den letzten Jahrzehnten durch eine Myriade von präklinischen Studien akkumuliert wurde, repräsentiert. Das Einpflegen von Konzepten in solch eine Datenbank setzt voraus, dass die Konzepte und Relationen der SCIO mit den korrespondierenden Begriffen aus der natürlichen Sprache, wie sie in wissenschaftlichen Publikationen vorkommen, vernetzt werden können. Die relationale Struktur und die daraus hervorgehende, gesamtheitliche Repräsentation einer präklinischen Studie und die überdurchschnittlich hohe Übereinstimmung beim Annotationsprojekt zeigen, dass dieses Vorhaben mithilfe der SCIO unterstützt werden kann. Offensichtlich kann durch die semiautomatische Datenextraktion mit SCIO das aktuell vorhandene Wissen aus präklinischen Studien

nicht vollständig erfasst werden. Die Ontologie kann zudem als Grundlage für automatisierte und skalierbare Prozesse dienen, die eine automatische Datenextraktion ermöglichen. So wurden im Rahmen dieser Dissertation weitere 50 Publikationen mit der SCIO annotiert. Paasen et al. [25] hat die relationale Struktur der SCIO und die mit ihr annotierten Konzepte und daraus resultierenden Datenpunkte aus 50 Publikationen als Grundlage für die Konzeption eines Modells zur automatischen Datenextraktion aus wissenschaftlichen Publikationen über präklinische Studien in der SCI-Domäne erfolgreich verwendet. Die reichhaltige relationale Struktur, die der SCIO zugrunde liegt, würde eine Ausweitung solch eines Datenextraktionsmodells ermöglichen, um im weiteren Verlauf größere, für die Ausführung von Metastudien notwendige, Wissenseinheiten einzufangen.

4.3 Schlussfolgerungen

Im Rahmen dieser Dissertation wurde die SCIO zur systematischen Erfassung des vorhandenen Wissens aus präklinischen Tierstudien über traumatische Rückenmarksverletzungen konzipiert. Die Ontologie wurde in einem Bottom-Up-Verfahren entwickelt, indem relevante Konzepte aus über 140 Publikationen manuell extrahiert und in sie integriert wurden. Die daraus entstandene Ontologie besteht aus 385 Klassen und 46 Eigenschaften und ist in der Lage eine Publikation aus der Domäne der SCI in Bezug auf ihren Inhalt und ihrer Signifikanz vollständig und präzise zu repräsentieren, damit das aus ihr hervorgehende Wissen für Wissenschaftler einfacher zugänglich und auffindbar ist. Darüber hinaus kann die Ontologie Metastudien und Entscheidungsfindungen für die Translation von geeigneten präklinischen Studien in klinische Studien, basierend auf der aktuell vorliegenden Evidenz, unterstützen [5]. Wie bereits im Mengendiagramm in Abb. 1 dargestellt, bietet die Architektur der SCIO eine Grundlage für die automatische Datenextraktion und für die Konzeption einer Datenbank. Basierend auf der im Rahmen dieser Dissertation entwickelten Ontologie und ihrer Anwendung bei der semi-automatischen Datenextraktion, wurde das PSINK-Project gegründet. Es besteht aus einer weiterentwickelten, online-verfügbaren Version der SCIO, deren mittels automatischer Datenextraktion aus wissenschaftlichen Texten extrahierte Konzepte aus präklinischen Studien über SCI in eine Onlinedatenbank integriert werden, um den aktuellen Wissensstand der SCI-Domäne vollständig abzubilden und frei zugänglich zu machen [5]. Darüber hinaus liefert die Struktur der

SCIO die Grundlagen für die Konzeption des SCIExplorers, einer Internetapplikation zur schnellen und zielgerichteten Abfrage und Lokalisation von präklinischen Forschungsdaten aus der PSINK-Datenbank [4]. Somit konnte die ursprüngliche Zielsetzung dieser Dissertation durch die Erweiterbarkeit der SCIO für die automatische Datenextraktion im Zusammenspiel mit einer frei zugänglichen Datenbank im PSINK-Projekt realisiert werden.

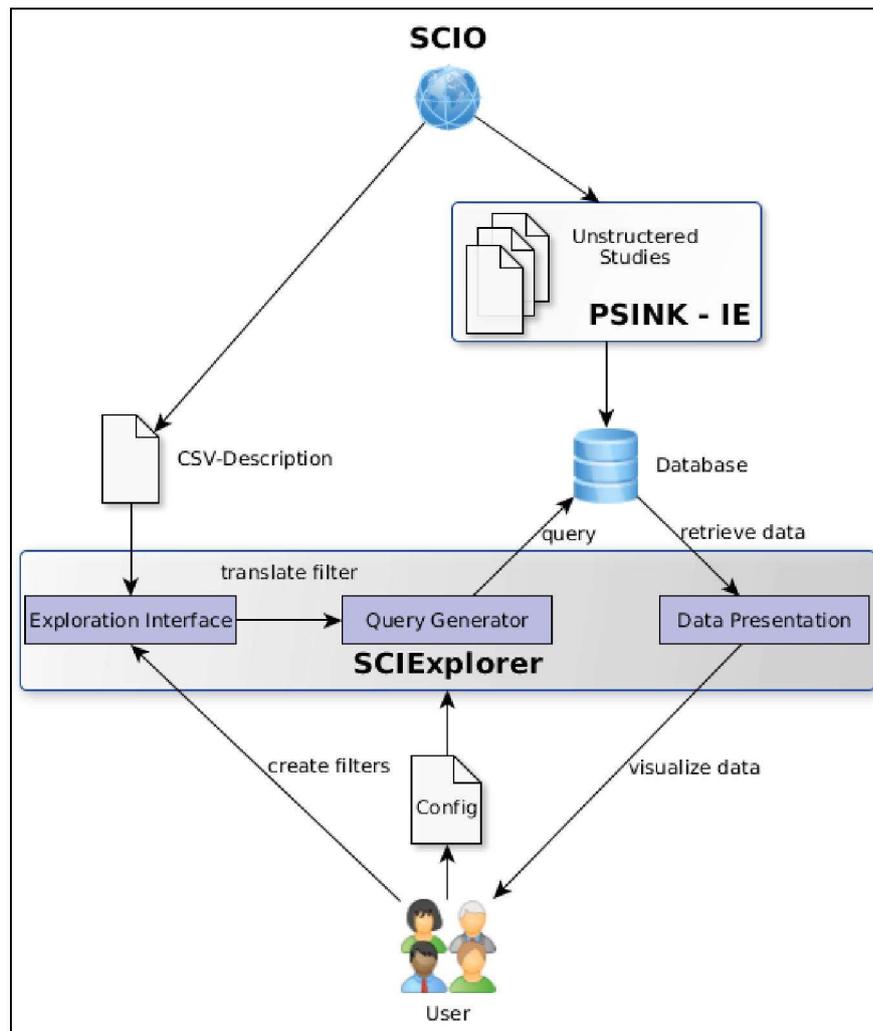


Abbildung 34: Borowi, Ter Horst et al.jpg [4] **Das PSINK-Projekt im Überblick.**

Anwender können mittels des SCIExplorers eine onlineverfügbare Datenbank durchsuchen. Die in die Datenbank eingepflegten Konzepte werden zuvor mit einer weiterentwickelten, onlineverfügbaren Version der SCIO durch automatische Datenextraktion aus wissenschaftlichen Texten gewonnen. Das aktuelle Domänenwissen soll somit frei zugänglich werden und fördert einen offenen Wissensaustausch für die Entwicklung erfolgsversprechender Therapien zur Behandlung von Patienten bei Zustand nach TRV.

Es kann weitgehend festgestellt werden, dass die in dieser Doktorarbeit erstellte Ontologie und ihre Applikation auf wissenschaftliche Texte in Kombination mit den Annotationsrichtlinien die Grundlagen für den Entwurf von weiteren Werkzeugen zur Förderung einer systematischen Repräsentation des vorhandenen Domänenwissens und eines offenen Wissensaustausches darstellt. Die hohen Cohen's Kappa Koeffizienten im oben beschriebenen Annotationsprojekt zeigen, dass die SCIO die präklinische Domäne der traumatischen Rückenmarksverletzungen vollständig repräsentiert und einen präzisen Einblick in die aktuelle Studienlage erlaubt. Zukünftige Einflussbereiche der SCIO können Richtlinien für die Erstellung von wissenschaftlichen Veröffentlichungen sein, die ein Mindestmaß an zu erwähnenden Konzepten voraussetzen, damit der experimentelle Aufbau, der experimentelle Ablauf und die Ergebnisse transparenter dargestellt werden, um eine angemessene Datenextraktion und einen daraus resultierenden Datenaustausch zu erleichtern. Weiterhin kann somit eine adäquate Reproduzierbarkeit von präklinischen Studien ermöglicht werden.

Die SCIO und die Annotationsrichtlinien sind auf der Homepage des CNR mit den folgenden Links verfügbar:

<http://www.cnr.de/files/upload/projekte/SCIO4.25.zip>

<http://www.cnr.de/files/upload/projekte/SCIO%20annotation%20guideline%20final%200draft.pdf>

4.4 Einfluss der SCIO auf die gute wissenschaftliche Praxis

Zur Verbesserung und Weiterentwicklung von experimentellen Studien werden Appelle von Wissenschaftlern an Förderinstitutionen immer lauter, die vollständige Darlegung aller experimentellen Daten als Grundvoraussetzung für den Erhalt von öffentlichen Fördergeldern einzufordern, denn die selektive Berichterstattung von lediglich positiven Resultaten kann weitreichende Folgen für das Wohlergehen von Patienten haben und zu einer Verschwendung von finanziellen Ressourcen führen, da vermeintlich negative Resultate ebenfalls wertvolle Anhaltspunkte für die Verbesserung von Therapieansätzen bedeuten können [7]. Dieser Aufruf ist verständlich und darauf zurückzuführen, dass in der Regel nicht alle Ergebnisse einer Studie in einer Publikation offengelegt werden, sondern nur positive Ergebnisse erwähnt werden, die von den Autoren und ausführenden Arbeitsgruppen als erfolgreiche Lösung zur Problemstellung gewertet

werden [27]. Negative Ergebnisse werden nur selten erwähnt oder diskutiert. Dabei ist in Betracht zu ziehen, dass die Extraktion von negativen Ergebnissen mithilfe der SCIO die Entwicklung von zielgerichteten Herangehensweisen bei der Identifikation von erfolgsversprechenden Therapien fördern kann, indem sie einen konstruktiven Wissensaustausch zwischen Arbeitsgruppen unterstützt. Somit kann der Inhalt geprüft und ggf. verbessert werden. Erst die Präsentation des gesamten Wissens erlaubt es Rückschlüsse und Verbesserungen beim experimentellen Aufbau vorzunehmen oder andere Lösungsansätze in Betracht zu ziehen, denn auch die Publikation von negativen Ergebnissen kann zur Entwicklung von evidenzbasierten Therapien beitragen [18].

Jegliche Form der Vorenthaltung von studienrelevanten Daten reduziert die Transparenz und erschwert die Reproduzierbarkeit von präklinischen Studien. Als Antwort auf die schwierige Reproduzierbarkeit von präklinischen Studien hat die NIH Trainingsmodule zur Konzeption eines guten experimentellen Designs, für eine Verbesserung der Transparenz und für eine eindeutige Darstellung von Ergebnissen vorgeschlagen [8]. Eine Konferenz des European College of Neuropsychopharmacology (ECNP) in Berlin, konstatierte ebenfalls, dass neben der Reproduzierbarkeit ein Training von Wissenschaftlern zur Aufrechterhaltung von guter wissenschaftlicher Praxis notwendig ist [31]. Maßnahmen dieser Art können von der ganzheitlichen Wissensextraktion mit SCIO profitieren, vorausgesetzt, dass das Ausmaß des in präklinischen Studien präsentierten Wissens vollständig ist und keine wichtigen Daten vorenthalten werden. Ein weiterer Aspekt der vollständigen Wissenspräsentation ist die sogenannte *long-tail-data*. Hierbei handelt es sich um unstrukturierte Datenvolumen, die im Gegensatz zur gut strukturierten, durch homogene Datenvolumen gekennzeichneten *big science*, aus nicht publizierter *dark data*, bestehen und lediglich einen kleinen Anteil des gesamten wissenschaftlichen Datenvolumens einnehmen, während der größte Anteil durch kleine unabhängige, individuelle Arbeitsgruppen generiert und in heterogener Form dargestellt wird [10]. Die Aufgabe der SCIO ist es das bereits vorhandene, heterogene Wissen durch eine gezielte Datenextraktion eindeutig und klar zu präsentieren und in Wissenschaftskreisen darauf aufmerksam zu machen. Die Erstellung der SCIO basiert auf der Formalisierung von Konzepten, die das vermeintlich strukturierte homogene Wissen aus wissenschaftlichen Publikationen repräsentiert. Somit verwendet sie den bereits bestehenden wissenschaftlichen Konsens bei der Zusammenfassung des extrahierten experimentellen Aufbaus, des experimentellen Ablaufs und der Ergebnisse

von wissenschaftlichen Studien. Auch *long-tail-data* birgt potenziell wertvolles Wissen, das für eine Weiterentwicklung von Therapien sehr hilfreich sein kann, aber nicht für die Wissensgemeinschaft zugänglich ist. Die Konzeptualisierung von unstrukturiertem und z.T. nicht präsentem Wissen ist mit der Ontologie nicht möglich. Sie ist zwangsläufig auf die Veröffentlichung von Studiendaten angewiesen und könnte das verborgene Wissen erst dann extrahieren, wenn Wissensdatenbanken zur Repräsentation der *long-tail-data* erstellt werden. Im Allgemeinen sollte ein gutes Studiendesign berücksichtigen, dass eine hohe Rate an Übereinstimmung bei der Beurteilung einer Studie durch unterschiedliche Personen gewährleistet werden muss [22]. Dieses Vorhaben kann durch ein eindeutiges und strukturiertes Berichterstattungsformat realisiert werden. Verbesserungen bei der Ausbildung von Wissenschaftlern, z.B. im Umgang mit experimentellen Materialien oder bei der Konzeption eines adäquaten experimentellen Aufbaus, tragen zur Verbesserung der Reproduzierbarkeit von biomedizinischen präklinischen Studien bei [11]. Durch die jährlich stetig wachsende Anzahl an Publikationen ist die Entwicklung der Ontologie ein dynamischer Prozess. [1] Neue experimentelle Methoden und andere Erkenntnisse, die aus den zahlreichen Publikationen hervorgehen, führen zu neuen Konzepten. Diese müssen unter Berücksichtigung ihrer Attribute in das Hauptklassen- und Unterklassengefüge der Ontologie integriert werden, um eine aktuelle Wissensrepräsentation der SCIO aufrechtzuerhalten. Unterstützend für dieses Vorhaben ist der Vorschlag, dass die Transparenz präklinischer Studien nur durch die Zusammenarbeit mehrerer Interessengruppen bestehend aus Forschern, Gutachtern, Fördereinrichtungen und Verlage erfolgen sollte [18]. Die SCIO bietet bereits sehr gute Voraussetzungen für eine adäquate Transparenz und fachgerechte Beurteilung durch die zuvor genannten Interessengruppen. Darüber hinaus kann ihre Weiterentwicklung durch genau diese Interessengruppen gefördert werden, um robuste Therapien für querschnittsgelähmte Personen zu entdecken.

Mein aufrichtiger Dank gilt Herrn Prof. Dr. Hans Werner Müller und Frau Dr. Nicole Brazda, die mich während der Bearbeitung meiner Dissertation durch einen offenen und zielorientierten fachlichen Austausch unterstützt haben.

Mein besonderer Dank gilt meiner Familie Jürgen Kirchhoffer †, Berlanti Kirchhoffer, Selim Kirchhoffer und meiner Ehefrau Ann Kirchhoffer, die mir während meiner akademischen Ausbildung und in allen wichtigen Abschnitten meines Lebens mit großer Unterstützung zur Seite standen.

References

- [1] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. 2000. Gene ontology: tool for the unification of biology. *Nature genetics* 25, 1, 25–29.
- [2] Baker, M. 2015. PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science (New York, N.Y.)* 349, 6251, aac4716.
- [3] Biering-Sørensen, F., Alai, S., Anderson, K., Charlifue, S., Chen, Y., DeVivo, M., Flanders, A. E., Jones, L., Kleitman, N., Lans, A., Noonan, V. K., Odenkirchen, J., Steeves, J., Tansey, K., Widerström-Noga, E., and Jakeman, L. B. 2015. Common Data Elements for Spinal Cord Injury Clinical Research: A National Institute for Neurological Disorders and Stroke Project. *Spinal cord* 53, 4, 1–33.
- [4] Borowi, A., Ter Horst, H., Hartung, M., Estrada, V., Brazda, N., Müller, H. W., and Cimiano, P., Eds. 2017. *Ontology-driven Visual Exploration of Preclinical Research Data in the Spinal Cord Injury Domain*. Proceedings of the SEMANTICS 2017 Poster and Demo Track 2044.
- [5] Brazda, N., Ter Horst, H., Hartung, M., Wiljes, C., Estrada, V., Klinger, R., Kuchinke, W., Müller, H. W., and Cimiano, P., Eds. 2017. *SCIO: An Ontology to Support the Formalization of Pre-Clinical Spinal Cord Injury Experiments*. Proceedings of the Joint Ontology Workshops 2017; Episode 3: The Tyrolean Autumn of Ontology 2050. CEUR Workshop Proceedings.
- [6] Carpentier, M., Combescure, C., Merlini, L., and Perneger, T. V. 2017. Kappa statistic to measure agreement beyond chance in free-response assessments. *BMC medical research methodology* 17, 1, 1–8.
- [7] Chan, A.-W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gøtzsche, P. C., Krumholz, H. M., Ghersi, D., and van der Worp, H. B. 2014. Increasing value and reducing waste: addressing inaccessible research. *The Lancet* 383, 9913, 1–25.
- [8] Collins, F. S. and Tabak, L. A. 2014. Policy: NIH plans to enhance reproducibility. *Nature* 505, 7485, 612–613.
- [9] Fehlings, M., Singh, A., Tetreault, L., Kalsi-Ryan, S., and Nouri, A. 2014. Global prevalence and incidence of traumatic spinal cord injury. *CLEP*, 309.
- [10] Ferguson, A. R., Nielson, J. L., Cragin, M. H., Bandrowski, A. E., and Martone, M. E. 2014. Big data from small data: data-sharing in the 'long tail' of neuroscience. *Nature neuroscience* 17, 11, 1442–1447.
- [11] Freedman, L. P., Cockburn, I. M., and Simcoe, T. S. 2015. The Economics of Reproducibility in Preclinical Research. *PLoS biology* 13, 6, e1002165.
- [12] Gašević, D., Djuric, D., and Devedžić, V. 2006. *Model Driven Architecture and Ontology Development*. Springer Berlin Heidelberg.
- [13] Gomes-Osman, J., Cortes, M., Guest, J., and Pascual-Leone, A. 2016. A Systematic Review of Experimental Strategies Aimed at Improving Motor Function after Acute and Chronic Spinal Cord Injury. *Journal of neurotrauma* 33, 5, 425–438.
- [14] Grüninger, M. and Fox, M. S., Eds. 1995. *Methodology for the Design and Evaluation of Ontologies*.
- [15] Karp, N. A. 2018. Reproducible preclinical research-Is embracing variability the answer? *PLoS biology* 16, 3, e2005413.

- [16] Kwon, B. K., Okon, E. B., Tsai, E., Beattie, M. S., Bresnahan, J. C., Magnuson, D. K., Reier, P. J., McTigue, D. M., Popovich, P. G., Blight, A. R., Oudega, M., Guest, J. D., Weaver, L. C., Fehlings, M. G., and Tetzlaff, W. 2011. A grading System To Evaluate Objectively the Strength of Pre-Clinical Data of Acute Neuroprotective Therapies for Clinical Translation in Spinal Cord Injury. *Journal of neurotrauma* 28, 8, 1525–1543.
- [17] Landis, J. R. and Koch, G. G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1, 159–174.
- [18] Landis, S. C., Amara, S. G., Asadullah, K., Austin, C. P., Blumenstein, R., Bradley, E. W., Crystal, R. G., Darnell, R. B., Ferrante, R. J., Fillit, H., Finkelstein, R., Fisher, M., Gendelman, H. E., Golub, R. M., Goudreau, J. L., Gross, R. A., Gubitza, A. K., Hesterlee, S. E., Howells, D. W., Huguenard, J., Kelner, K., Koroshetz, W., Krainc, D., Lazic, S. E., Levine, M. S., Macleod, M. R., McCall, J. M., Moxley, R. T., Narasimhan, K., Noble, L. J., Perrin, S., Porter, J. D., Steward, O., Unger, E., Utz, U., and Silberberg, S. D. 2012. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490, 7419, 1–10.
- [19] Lemmon, V. P., Abeyruwan, S., Visser, U., and Bixby, J. L. 2014. Facilitating transparency in spinal cord injury studies using data standards and ontologies. *Neural regeneration research* 9, 1, 6–7.
- [20] Lemmon, V. P., Ferguson, A. R., Popovich, P. G., Xu, X.-M., Snow, D. M., Igarashi, M., Beattie, C. E., and Bixby, J. L. 2014. Minimum Information about a Spinal Cord Injury Experiment: A Proposed Reporting Standard for Spinal Cord Injury Experiments. *Journal of neurotrauma* 31, 15, 1354–1361.
- [21] Mariano Fernandez and Asuncion Gomez-Pearez and Natalia Juristo. Methontology: From Ontological Art Towards Ontological Engineering.
- [22] McHugh, M. L. 2012. Interrater reliability: the kappa statistic. *Biochem Med*, 276–282.
- [23] Nielson, J. L., Guandique, C. F., Liu, A. W., Burke, D. A., Lash, A. T., Moseanko, R., Hawbecker, S., Strand, S. C., Zdunowski, S., Irvine, K.-A., Brock, J. H., Nout-Lomas, Y. S., Gensel, J. C., Anderson, K. D., Segal, M. R., Rosenzweig, E. S., Magnuson, D. S. K., Whittemore, S. R., McTigue, D. M., Popovich, P. G., Rabchevsky, A. G., Scheff, S. W., Steward, O., Courtine, G., Edgerton, V. R., Tuszynski, M. H., Beattie, M. S., Bresnahan, J. C., and Ferguson, A. R. 2014. Development of a Database for Translational Spinal Cord Injury Research. *Journal of neurotrauma* 31, 21, 1789–1799.
- [24] NIST Big Data Public Working Group. 2015. *NIST Big Data Interoperability Framework: Volume 1, Definitions*. National Institute of Standards and Technology.
- [25] Paassen, B., Stöckel, A., Dickfelder, R., Göpfert, J. P., Brazda, N., Kirchhoffer, T., Müller, H. W., Klinger, R., Hartung, M., and Cimiano, P., Eds. 2014. *Ontology-based Extraction of Structured Information from Publications on Preclinical Experiments for Spinal Cord Injury Treatments*. Association for Computational Linguistics and Dublin City University, Dublin, Ireland.
- [26] Pedro-Roig, L. and Emmerich, C. H. 2017. The reproducibility crisis in preclinical research – lessons to learn from clinical research. *MEW* 26, 28–32.

- [27] Prinz, F., Schlange, T., and Asadullah, K. 2011. Believe it or not: how much can we rely on published data on potential drug targets? *Nature reviews. Drug discovery* 10, 9, 328–329.
- [28] Ramer, L. M., Ramer, M. S., and Bradbury, E. J. 2014. Restoring function after spinal cord injury: towards clinical translation of experimental strategies. *The Lancet. Neurology* 13, 12, 1241–1256.
- [29] Sejnowski, T. J., Churchland, P. S., and Movshon, J. A. 2014. Putting big data to good use in neuroscience. *Nature neuroscience* 17, 11, 1440–1441.
- [30] Silva, N. A., Sousa, N., Reis, R. L., and Salgado, A. J. 2014. From basics to clinical: a comprehensive review on spinal cord injury. *Progress in neurobiology* 114, 25–57.
- [31] Steckler, T., Brose, K., Haas, M., Kas, M. J., Koustova, E., and Bernalov, A. 2015. The preclinical data forum network: A new ECNP initiative to improve data quality and robustness for (preclinical) neuroscience. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 25, 10, 1803–1807.
- [32] Xie, Z., Gadepalli, C., and Cheetham, B. M.G. 2017. REFORMULATION AND GENERALISATION OF THE COHEN AND FLEISS KAPPAS. *lijhls* 3, 3, 1–15.