SCIO: An Ontology to Support the Formalization of Pre-Clinical Spinal Cord Injury Experiments

Nicole BRAZDA [b], Hendrik TER HORST [a], Matthias HARTUNG [a], Cord WILJES [a], Veronica ESTRADA [b], Roman KLINGER [c], Wolfgang KUCHINKE [d], Hans Werner MÜLLER [b], Philipp CIMIANO [a]

[a] *Exzellenzcluster Kognitive Interaktionstechnologie (CITEC), Bielefeld University, Germany*
[b] *Center for Neuronal Regeneration (CNR) and Neurology, Heinrich Heine University Düsseldorf, Germany*
[c] *Institut für Maschinelle Sprachverarbeitung (IMS), University of Stuttgart, Germany*
[d] *Coordination Center for Clinical Trials, Heinrich Heine University Düsseldorf, Germany*

Abstract. We present the Spinal Cord Injury Ontology (SCIO), which has the goal to support the representation of pre-clinical studies in the domain of spinal cord injury (SCI) therapies. The ontology is developed in the context of the PSINK project, as part of an information extraction lifecycle to populate a database with comprehensive knowledge about pre-clinical studies published in the SCI literature. This database enables domain experts to explore and access all relevant knowledge that is available as a basis to support clinical decision-making and translation of pre-clinical evidence into clinical practice. Here, we discuss the methodology underlying the development of SCIO and the main design choices made throughout. We also present a web application that relies on SCIO to organize pre-clinical knowledge and present it to domain experts for exploration purposes. In particular, the application enables experts to get relevant answers and insights on the outcomes of different therapies in pre-clinical studies and how their effectiveness varies depending on core parameters such as injury type, dosage, time of application, or investigation method applied, among others.

Keywords. Spinal Cord Injury, Ontology, Pre-clinical Studies, Information Extraction

1. Introduction

On every single day, thousands of new scientific publications are uploaded to PubMed[1], the standard repository for biomedical literature. The subset of literature that is relevant for spinal cord injury treatment comprises more than 130,000 publications currently

[1] https://www.ncbi.nlm.nih.gov/pubmed

available in PubMed, among them approx. 18,000 pre-clinical SCI studies.[2] It is impossible for any research group, let alone individual researchers, to keep up with the amount of medical knowledge published in their specific area. Yet, taking decisions in the interest of promoting the state of the art or translating the available evidence from pre-clinical or clinical studies into therapeutic concepts requires knowledge of all, or at least the most important studies available. Meta-studies or systematic reviews obviously can help in this respect, as they provide an overview of the most important results in the limited number of papers included in such a meta-analysis. However, the creation of systematic reviews is time-consuming. Thus, by the publication time of the review, some of the reviewed results may already be outdated.

The goal of the PSINK project[3] is to develop an information extraction workflow for automatically gathering the main parameters of pre-clinical studies and use them to populate a database that supports the exploration of the entirety of available pre-clinical evidence. In PSINK, we are particularly concerned with aggregating available pre-clinical evidence to support decision-making on which therapies might be prospective candidates to develop a successful therapy to cure spinal cord injuries in human patients, thus fostering the translation of pre-clinical therapeutic concepts into clinical practice.

For this purpose, the presented ontology supports the formalization of the structure, parameters and results of a pre-clinical study. While the ontology has been designed with the purpose of capturing pre-clinical studies in the spinal cord injury domain, its core can be used to represent pre-clinical (and, partially, clinical) trials in other medical domains as well. In this paper, we present the design of the ontology, the top-level structure and which ontologies have been reused. We further describe our efforts to align SCIO with other existing vocabularies. As a proof of concept, we present a web application that has been developed based on the ontology and enables experts to explore the available evidence.

2. Ontology Design

The Spinal Cord Injury Ontology has been designed following the methodology developed in the On-To-Knowledge project [1] and described by Sure et al. [2]. Domain experts in the field of spinal cord injury with substantial experience in performing pre-clinical studies have been involved in all steps of the process.

The methodology comprises the four steps (i) *Kick-off* (definition of competency questions that the ontology is expected to answer); (ii) *Refinement* (implementation in the Web Ontology language using Protégé); (iii) *Evaluation* with respect to the competency questions defined in the kick-off phase; (iv) *Application and Evolution*. In the latter phase, the ontology has been used to automatically derive an annotation scheme to be used to annotate a number of articles that the information extraction system can be trained on. Moreover, the ontology has been used as the conceptual backbone for developing a web application that supports the exploration of the available evidence (*cf.* Section 3).

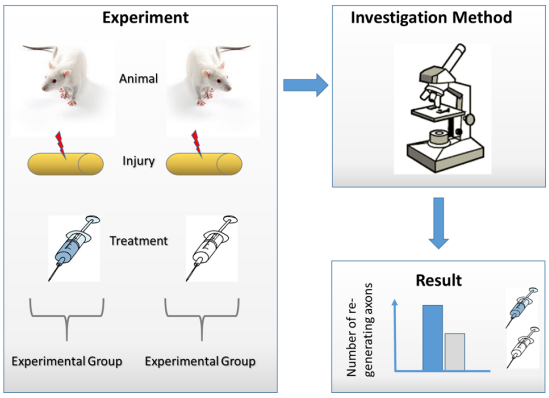

Figure 1. Schematic overview of the basic logic in SCI pre-clinical experiments.

2.1. Scope and Competency Questions

SCIO has been designed to mirror the standard workflow and basic configuration of a pre-clinical study in the SCI domain: An animal model is selected and groups of animals are defined which receive a certain type of injury and a treatment or no treatment, before being examined for recovery. A schematic overview of this fundamental structure in SCI pre-clinical experiments is shown in Figure 1. Even if multiple experimental groups are compared to each other, the direct comparison and statistical analysis is based on two groups with a defined setting of animals, experimental spinal cord injury and treatment. In the example depicted in Figure 1, the two groups differ in treatment type. This could, *e.g.*, be subcutaneous application of a drug vs. application of the buffer as control.

The investigation method can be a histological analysis of the spinal cord tissue, a molecular analysis or a functional/behavioural test (*e.g.*, a horizontal ladder walking test [3]), all of which are included in SCIO. The result of a test is either a between-group difference of an investigated effect, *e.g.*, the number of regenerating axons, or no change.

In the kick-off phase, the design team has defined a number of exemplary competency questions that the ontology should be able to answer:

(1) Which treatments yielded positive results in different animal models?
(2) Which treatments yielded positive results in functional as well as in non-functional tests?
(3) Which treatments show positive results only in partial lesions but not in complete lesions?
(4) Which treatments show a functional effect for thoracic as well as cervical lesions?
(5) In which lesion models being tested in male rats or mice have no negative effects of erythropoietin been observed so far?
(6) After how many weeks can functional improvements of severe thoracic contusions be expected?
(7) Which investigation methods reveal the earliest differences in functional or non-functional tests between treatment and control groups?

[2]Results from querying PubMed for spinal cord injury and traumatic brain injury, as of September 13, 2017.
[3]`http://www.psink.de`

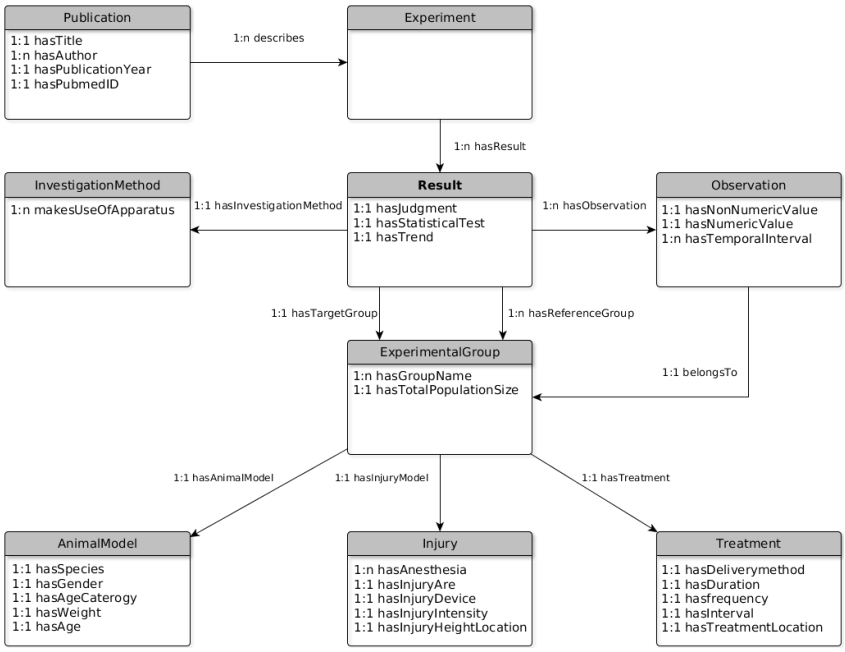

Figure 2. Ontological architecture of the main classes in SCIO.

2.2. Main Classes and Properties

Figure 2 shows the relations between the top-level classes in SCIO: *Publication*, *Experiment*, *Result*, *Observation*, *InvestigationMethod*, *ExperimentalGroup*, *OrganismModel*, *Injury*, and *Treatment*. The relations between these classes are as follows: Each *Publication* describes one or more experiments. Each *Experiment* consists of one or more results. Each *Result* is related to exactly one *InvestigationMethod*, one *TargetGroup* (typically the treated group), and a *ReferenceGroup* (*e.g.*, a control group). A *Result* consists of a number of *Observations* that are specific for one of the investigated experimental groups. An *ExperimentalGroup* is defined by exactly one *AnimalModel*, one *InjuryModel* and one *Treatment*. Overall, SCIO contains more than 500 manually added classes and 80 properties (data type and object type properties). We describe these classes in more detail below, giving specific examples of RDF code that model a complete study.

Publication: Each publication is described by the meta-data associated with the respective paper. This includes a unique identifier (*e.g.*, PubMed ID) to ensure provenance tracking, a list of authors, the year of publication etc. A publication describes one or more experiments, each of which is related to one or more results. The RDF code[4] in Figure 3 shows an example.

Result: A result describes the outcome of a test within a specific setting, based on a comparison of a reference group to a target group (*cf. ExperimentalGroup*). The setting is mainly determined by an investigation method (*cf. InvestigationMethod*) and a statistical

[4]All examples in this paper are given in RDF Turtle format, using `@prefix scio: <http://psink.de/scio/>` as a prefix referring to the SCIO namespace.

```
<scio:data/Publication_8>
        a       <scio:Publication> ;
        <scio:describes>
                <scio:data/Experiment_8> ;
        <scio:hasAuthor>
                <scio:data/Person_8> ;
        <scio:hasPublicationYear>
                "2009" ;
        <scio:hasPubmedID>
                "19372269" .
```

Figure 3. RDF excerpt describing a *Publication* instance.

```
<scio:data/Result_526>
        a       <scio:Result> ;
        <scio:hasInvestigationMethod>
                <scio:data/BBBTest_10> ;
        <scio:hasJudgement>
                <scio:Positive> ;
        <scio:hasObservation>
                <scio:data/Observation_1052> ,
                <scio:data/Observation_1053> ;
        <scio:hasReferenceGroup>
                <scio:data/DefinedExperimentalGroup_1052> ;
        <scio:hasStatisticalTest>
                <scio:data/StatisticalTest_0> ;
        <scio:hasTargetGroup>
                <scio:data/DefinedExperimentalGroup_1053> ;
        <scio:hasTrend>
                <scio:data/Decrease_230> .
```

Figure 4. RDF excerpt describing a *Result* instance.

test. The subjective interpretation of the result is modelled via the (nominal) class *Judgement*. Each result contains arbitrary observations supporting the author's judgement. The statistical observations describe the measurable *Trend* of the investigation method based on the used statistical test (*e.g.*, t-test).[5] The RDF code in Figure 4 shows a result obtained using a *BBBTest* as investigation method which yields two observations (one for each experimental group), and for which a positive judgement can be derived based on one statistical test. The Trend 'Decrease' represents the fact that the observed score of the BBB test [4] is lower in the target group than in the reference group.

Observation: An observation represents a quantitative or qualitative measurement conducted for a specific *ExperimentalGroup* (reference or target). Being related to a particular timepoint, each *Observation* is modelled as a *Perdurant* or *Ocurrent*. The observation may store numeric values (*e.g.*, a BBB score within a locomotor test) or non-numeric values in case of vague descriptions such as "higher, weaker, better…". Figure 5 shows

[5]Note that a decreasing trend does not strictly imply a negative judgement, as the trend is an *objective* observation, whereas the judgement is a *subjective* opinion based on the investigation method and other variables.

```
<scio:data/Observation_1053>
        a       <scio:Observation> ;
        <scio:belongsTo>
                <scio:data/DefinedExperimentalGroup_1053> ;
        <scio:hasNumericValue>
                "12,5" ;
        <scio:hasTemporalInterval>
                <scio:data/TemporalInterval_526> .
```

Figure 5. RDF excerpt describing an *Observation* instance.

```
<scio:data/DefinedExperimentalGroup_99>
        a       <scio:DefinedExperimentalGroup> ;
        <scio:hasInjuryModel>
                <scio:data/Contusion_3> ;
        <scio:hasNNumber>
                "12" ;
        <scio:hasOrganismModel>
                <scio:data/RatModel_6> ;
        <scio:hasTreatmentType>
                <scio:data/CompoundTreatment_518> .
```

Figure 6. RDF excerpt describing an experimental group instantiated as *DefinedExperimentalGroup*.

RDF code modelling an observation conducted during a particular temporal interval in which a value of 12.5 was measured in a BBB test on the target group.

InvestigationMethod: A result is related to exactly one *InvestigationMethod* which leads to an observation. Depending on the particular investigation method, different properties are used to specify its characteristics.

ExperimentalGroup: The *ExperimentalGroup* describes a set of animals (*cf. OrganismModel*) that were injured by a specific type of lesion (including sham injuries; *cf. InjuryType*) and subsequently received a treatment (including sham treatments; *cf. Treatment*). We distinguish two types of experimental groups: While a *DefinedExperimentalGroup* is explicitly defined by the author, an *AnalyzedExperimentalGroup* may refer to a sub-group and/or pooling of experimental groups. This enables arbitrary aggregation of groups for analysis. Thus, the ontology is not limited to compare only two groups (target vs. control group), but an arbitrary number of groups being treated or injured differently. This could be the case if the author clusters multiple experimental groups that received the same substance but with different dosages. Figure 6 shows an example.

OrganismModel: The *OrganismModel* describes the animal model that was used in the experiments together with its properties. An organism model is defined by its species, gender, weight, and age. We distinguish between categorical ages, i.e., “adult” or “young”, and non-categorical ages, *e.g.*, “3 months”. The RDF code in Figure 7 describes a rat model consisting of male, adult rats of the subspecies *SpragueDawleyRat* and weighing 312.5g (on average).

```
<scio:data/RatModel_6>
        a       <scio:RatModel> ;
        <scio:hasAgeCategory>
                <scio:Adult> ;
        <scio:hasGender>
                <scio:Male> ;
        <scio:hasSpecies>
                <scio:SpragueDawleyRat> ;
        <scio:hasWeight>
                "312.5 g" .
```

Figure 7. RDF excerpt describing an adult, male Sprague Dawley Rat model.

```
<scio:data/Contusion_3>
        a       <scio:Contusion> ;
        <scio:hasInjuryDevice>
                <scio:data/NYUImpactor_73> ;
        <scio:hasInjuryLocation>
                <scio:Thoracic> .
```

Figure 8. RDF excerpt describing a contusion applied using an NYU Impactor at thoracic level.

Injury: An *Injury* represents a type of injury that was applied to the animals in the treatment group. A description of the *Injury* includes the device that was used to cause the spinal cord injury, the area and the height (location) of the injury. Besides those main properties, an injury type comprises information about pre- and post-medication, anesthesia, and further animal care conditions such as housing, nutrition, and hydration, which are all modelled using object properties. The RDF code in Figure 8, for instance, describes a *Contusion* applied via an *NYU Impactor* at *Thoracic* level.

Treatment: A *Treatment* represents the application of a drug (in case of a *CompoundTreatment*), device or other therapeutic intervention (*e.g.*, rehabilitative training) at a particular location of the spinal cord, with a specific dosage or a specific intensity, respectively. A treatment can be applied once, at intervals or for a specific duration. The RDF code in Figure 9, for instance, represents a compound treatment that is delivered intraperitoneally and has a specific *Dosage*. The *SuppliedCompound* stands for a compound produced by a certain supplier.

2.3. Subclasses in SCIO

For each of the top level classes described above, a number of subclasses have been introduced manually through observation in scientific publications: *InvestigationMethod* (88 subclasses), *InjuryType* (11), *InjuryDevice* (30), *Anaesthetics & Medication* (21), *AnimalSpecies* (17), *Treatment* (excl. *CompoundTreatment*; 17), *CompoundTreatment* (applied compounds; 84).

2.4. Alignment to other Ontologies

SCIO imports or is aligned with several ontologies:

```
<scio:data/CompoundTreatment_518>
        a       <scio:CompoundTreatment> ;
        <scio:hasDeliveryMethod>
                <scio:IntraperitonealDelivery> ;
        <scio:hasDosage>
                 <scio:data/Dosage_31>;
        <scio:hasSuppliedCompound>
                <scio:data/SuppliedCompound_18> .
```

Figure 9. RDF excerpt describing a *Treatment* instance.

```
<scio:data/TemporalInterval_445>
        a       <scio:TemporalInterval> ;
        <scio:hasTemporalDuration>
                <scio:data/TemporalDuration_4>
        <scio:hasEventAfter>
                <scio:data/CompoundTreatment_429> ;
        <scio:hasEventBefore>
                <scio:data/Contusion_185> .

<scio:data/TemporalDuration_4>
        a       time:Duration ;
        time:numericDuration
                "1"^^xsd:decimal ;
        time:#unitType
                time:unitWeek .
```

Figure 10. RDF code showing how to model the time passed between injury and treatment in terms of the duration of an intermediate interval.

Time Ontology: Modeling the temporal structure of events is a key component in capturing the core parameters of a pre-clinical trial. Thus, the entities *Injury*, *Treatment* and *Observation* are modelled in SCIO as perdurants or occurrents according to the W3C Time Ontology[6]. They are modelled as subclasses of `scio:Event`, which is a subclass of `time:TemporalEntity`. The time passed between an injury and treatment is modelled via an interval that immediately succeeds the injury interval and precedes the treatment interval, as shown in the RDF code in Figure 10. This example models the situation that a treatment is applied one week after injury in terms of an intermediate interval with the respective duration.

QUDT Ontology: We reuse the QUDT Ontology[7] to model quantities and their units. In particular, the following classes are modelled as subclasses of `qudt:Quantity`: *Temperature*, *Force*, *ElectricFieldStrength*, *Duration*, *MeanValue*, *Dosage*, *Pressure*, *Longitude*, *Weight*, *Depth*, *Thickness*, *MedianValue*, *Volume*, *DosageExtracorporal*, *Voltage*, *DosageIntracorporal*, *NumericValue*, *Current*, *Distance*, *Age*. The RDF code in Figure 11 shows how a dosage of 20,000 units per liter is modelled as an instance of *Quantity*.

[6] `http://www.w3.org/2006/time`
[7] `http://data.qudt.org/qudt/owl/1.0.0/`

```
<scio:data/Dosage_31>
        a       qudt:Quantity ;
        qudt:quantityValue [
        qudt:unit qudt#InternationalUnitPerLiter> ;
        qudt:numericValue> "20,000" ].
```

Figure 11. RDF code showing how a dosage as consisting of a unit and a value is modelled as an instance of Quantity.

```
?Result <rdf:type> <scio:Result>.
?Result <scio:hasTargetGroup> ?TargetExperimentalGroup.
?TargetExperimentalGroup <scio:hasOrganismModel> ?OrganismModel.
?OrganismModel <scio:hasGender> ?Gender.
?OrganismModel <scio:hasGender> <scio:Male>.
{?OrganismModel <rdf:type> <scio:RatModel>. }
UNION
{?OrganismModel <rdf:type> <scio:MouseModel>. }.
?TargetExperimentalGroup <scio:hasTreatmentType> ?TreatmentTypes.
?TreatmentTypes <scio:hasSuppliedCompound> ?SuppliedCompound.
?SuppliedCompound <scio:hasCompound> ?Compound.
?SuppliedCompound <scio:hasCompound> <scio:Erythropoietin>.
```

Figure 12. SPARQL query generated by SCIExplorer for answering competency question (5) in Section 2.1.

3. Proof of Concept

We have implemented the web application SCIExplorer[8] that enables domain experts to explore the available evidence and to answer questions regarding the positive or negative effect of a treatment under different side conditions formalized as filters based on SCIO concepts [5]. The data underlying SCIExplorer was manually gathered by domain experts[9] in a process of analyzing 140 scientific articles and entering their key parameters into a spreadsheet which was then automatically transferred to RDF using SCIO. When accessing the web application, users can enter a potential therapy and get an overview of different diagrams plotting the ratio of positive to negative results on a therapy over animal models.

As a proof of concept, we demonstrate how competency question (5) as introduced in Section 2.1 can be answered using SCIExplorer: Figure 12 shows the SPARQL query that has been automatically generated from the appropriate filter settings[10]. The resulting RDF triples matching this query can be retrieved from `http://psink.techfak.uni-bielefeld.de/scio/examples/CQ5-result.rdf`.

Analogously, competency questions (1)–(3) can be answered via SCIExplorer by accessing the relevant information for specific treatments[11]. For competency question

[8] `http://psink.techfak.uni-bielefeld.de/SCIExplorer/`

[9] In the PSINK project, we are developing a semi-automatic workflow for robust, large-scale knowledge extraction using text mining and information extraction methods.

[10] `https://tinyurl.com/SCIExplorer-settings`

[11] The view for Estrogen treatments, *e.g.*, can be found at `https://tinyurl.com/scioestrogentreatment`.

(4), the information can be retrieved by setting a filter on *InjuryType→InjuryVertebral-Location* to *Cervical* and *Thoracic*.

4. Related Work

Several ontologies have been proposed to formally capture the structure and outcomes of *pre-clinical* or *clinical studies*. A prominent example for an ontology for *clinical studies* is the PICO ontology, developed by the Cochrane Foundation [7]. The main entity in the PICO ontology is a *PICO* class to represent the main aspects of a clinical study: *Patient*, *Population*, *Problem*, *Intervention*, *Comparison* and *Outcome*. PICO proposes a star-based modelling scheme in which a number of object properties are connected to an instance of the PICO class, in particular: *population*, *excludedPopulation*, *intervention-Group* and *outcomeGroup*. *Populations* are described via *age* (with class *Age* as range), *condition* (with class *Condition* as range) and *sex* (with class *Sex* as range) properties. An *OutcomeGroup* (range of *outcomeGroup*) represents a group of outcomes. Corresponding classes in PICO and SCIO are *Population* and *AnimalModel*, as well as *Outcome* and *Result*. However, the modelling of the actual experimental result is much more detailed in SCIO than in PICO. Rather than modeling the *Result* as an atomic entity, SCIO captures the different experimental groups, the instrument and test applied to measure the outcome, the exact location of application and injury, the time of application, the result of statistical tests applied, etc. The modelling of the structure of a study is thus richer in the case of SCIO .

Khoo et al. developed an ontology to represent disease treatment information in medical abstracts [11]. In their modelling, an instance of a *Disease-Treatment* class is related to a *Condition*, a *Treatment*, a *Disease*, an *Effect* and *Evidence* for this effect. The *Disease-Treatment* class is thus equivalent to the class *PICO* in the PICO ontology. However, the modelling is more fine-grained than in PICO with administration schemes that include frequency and duration for Treatments. Measurements are modeled similarly to SCIO, whereas different types of effects are distinguished: disease effects, side effects and patient effects. The *Modality* class corresponds to the *Judgement* class in SCIO, representing the truth value of the occurrence of the effect. The *Condition* class has subclasses for *PatientCondition*, *TreatmentCondition* and *DiseaseCondition*. The *Evidence* class represents statistical evidence including sampling information as well as information about the experimental subjects. In contrast to SCIO, there is no detailed vocabulary to capture the statistical information in detail (*e.g.*, which statistical test, which p-value) nor vocabulary to capture details of the experimental subjects.

Our work is also related to the efforts of developing an ontology of clinical research, *viz.*, the OCRe ontology [9,10]. Significant differences to SCIO are: (i) OCRe was designed to represent human studies; (ii) its focus is on the representation of the study protocol, *i.e.*, the abstract representation of the scientific design of a clinical study. Representing actual results and observations produced in a study is out of the scope of OCRe.

Regarding *pre-clinical studies*, most closely related to our work is the RegenBase project [6], which also develops an ontology and knowledge base of SCI biology. The main publication of the project does not fully explain the main design choices. We there-

```
PREFIX hyque: <http://semanticscience.org/ontology/hyque.owl#>
PREFIX regenbase: <http://regenbase.org/ontology#>
PREFIX sio: <http://semanticscience.org/resource/>

# return distinct agent PubChem compound identifiers
SELECT DISTINCT ?pubchem_id
WHERE {
    # retrieve agent, target, and effect
    ?event hyque:HYPOTHESIS_0000015 ?agent .
    ?agent sio:SIO_000671 ?pubchem_id .
    ?event hyque:HYPOTHESIS_0000016 ?target .
    # filter agent to be a small molecule
    ?agent rdf:type regenbase:RB_0000125 .
    ?agent ?effect ?target .
    ?effect rdfs:label ?effect_label .
    # filter effect to be 'increase'
    FILTER(?effect_label = "increase") .
    # filter target to be an outcome measure or any subclass
    ?target rdf:type/rdfs:subClassOf* regenbase:RB_0008016 .
}
```

Figure 13. Example query for answering a competency question using the RegenBase Ontology [6]

fore compare our project to their online version of the ontology[12] based on our competency questions.

On the website of RegenBase, the authors provide the example query shown in Figure 13 that can be evaluated on the RegenBase SPARQL endpoint, answering the question: *What perturbagens have been observed to improve behavioral outcomes following injury?* As shown in Figure 13, RegenBase represents hypotheses using HYQUE [8], a system and vocabulary supporting hypothesis formulation and evaluation. In the above query, the variable `?agent` ranges over possible substances and is constrained to a small molecule (`regenbase:RB_0008016`). The variable `?target` ranges over entities defined in the RegenBase ontology and can be bound to different biological processes such as gene expression, a phosphorylation, a behavioural test. The `?target` variable is constrained to be of type 'behavioural assessment' (`regenbase:RB_0008016`). The relation between an `?agent` and a `?target` is modelled using a property instance to which information characterizing the relation can be added. The fact that the `?agent` improves or increases the `?target` biological process is expressed via an `rdfs:label` on the individual property and is constrained to the string 'increase'. Thus, neither is the relation between an `?agent` and a `?target` modelled from an ontological point of view, nor is there an ontologically appropriate way to represent improvements in RegenBase. According to these design choices, it is presumable that the potential future functions of RegenBase do not include modelling of more complex questions as targeted by SCIO, *e.g.*, regarding the effects of dosage, time point of application, or delivery method. Overall, RegenBase seems to be designed for different goals and might complement SCIO with respect to a mechanistic understanding of molecular pathways underlying study results in the future. Questions focussing on study design choices

[12] `https://bioportal.bioontology.org/ontologies/RB`

in pre-clinical experiments, however, can so far only be addressed by SCIO, since this is the only available ontology modelling the central relations in this respect.

5. Conclusion

This paper discusses SCIO, a novel ontology to formalize pre-clinical studies in the spinal cord injury domain. Its core structure is driven by design decisions to enable fine-grained representations of the specific characteristics of experiments in the domain.

The web application SCIExplorer serves as a proof of concept. It structures the pre-clinical evidence which is available and supports domain experts in the exploration of this information. Future work will focus on developing an information extraction pipeline in which the ontology represents a core element.

SCIO is available for download at `http://psink.de/scio/`.

Acknowledgments

This work has been funded by the German Ministry of Education and Research in the project PSINK (*Populating a Pre-clinical Spinal Cord Injury Knowledge Base to Support Clinical Translation*) under project numbers 031L0028A and 031L0028B.

References

[1] J. Davies, D. Fensel, and F. van Harmelen, editors. On-To-Knowledge. Semantic Web enabled Knowledge Management. Wiley (2002).

[2] York Sure, Steffen Staab, Rudi Studer: Ontology Engineering Methodology. In: Steffen Staab, Rudi Studer (eds.): Handbook on Ontologies. Springer (2009): 135–152.

[3] Gerlinde A. Metz and Ian Q. Whishaw: The ladder rung walking task. A scoring system and its practical application. J Vis Exp 28 (2009): 1204.

[4] D. Michele Basso, Michael S. Beattie, and Jacqueline C. Bresnahan: A sensitive and reliable locomotor rating scale for open field testing in rats. J Neurotrauma 12 (1995): 1–21.

[5] Alexander Borowi, Hendrik ter Horst, Matthias Hartung, Nicole Brazda, Veronica Estrada, Hans Werner Müller, Philipp Cimiano: Ontology-driven Visual Exploration of Pre-clinical Research Data in the Spinal Cord Injury Domain. Proc. of the SEMANTiCS Poster and Demo Track (2017). In print.

[6] A. Callahan, S.W. Abeyruwan, H. Al-Ali, K. Sakurai, A.R. Ferguson, P.G. Popovich, N.H. Shah, U. Visser, J.L. Bixby, and V.P. Lemmon: RegenBase. A Knowledge Base of Spinal Cord Injury Biology for Translational Research. Database (Oxford): pii: baw040 (2016).

[7] Chris Mavergames, Silver Oliver, and Lorne Becker: Systematic Reviews as an Interface to the Web of (Trial) Data. Using PICO as an Ontology for Knowledge Synthesis in Evidence-based Healthcare Research. Proc. of the Workshop on Semantic Publishing (2013): 22–26.

[8] Alison Callahan, Michel Dumontier, and Nigam H. Shah: HyQue. Evaluating Hypotheses Using Semantic Web Technologies. J Biomed Sem 2.Suppl 2 (2011): S3.

[9] Ida Sim, Samson W. Tu, Simona Carini, Harold P. Lehmann, Brad H. Pollock, Mor Peleg, Knut M. Wittkowski: The Ontology of Clinical Research (OCRe). An Informatics Foundation for the Science of Clinical Research. J Biomed Inform 52 (2013): 78–91.

[10] Samson W. Tu, Simona Carini, Alan Rector, Peter MacCallum, Igor Toujilov, Steve Harris, and Ida Sim: OCRe. An ontology of clinical research. Proc. of the 11th International Protégé Conference (2009).

[11] Christopher S.G. Khoo, Jin-Cheon Na, Vivian Wei Wang, and Syin Chan: Developing an Ontology for Encoding Disease Treatment Information in Medical Abstracts. DESIDOC Journal of Library & Information Technology 31 (2011): 103–115.