

A Systematic Survey of Temporal Requirements of Bio-Health Ontologies

Jared Leo [a], Nicolas Matentzoglu [a], Uli Sattler [a] and Bijan Parsia [a]
[a] *Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*
E-mails: jared.leo@manchester.ac.uk, nico.matentzoglu@manchester.ac.uk, uli.sattler@manchester.ac.uk, bijan.parsia@manchester.ac.uk

Abstract. The Description Logic $\mathcal{SROIQ}(\mathcal{D})$, as the logical core of the W3C standard Web Ontology Language (OWL 2), is a widely used formalism for ontologies in the life sciences. Bio-health applications including healthcare and life science domains commonly have a need to represent temporal information such as medication frequency or stage-based development. Different classes of temporal phenomena may generate different sorts of requirements on $\mathcal{SROIQ}(\mathcal{D})$ or extensions of $\mathcal{SROIQ}(\mathcal{D})$. In this paper, we deliver the first precise investigation into identifying exactly what kinds of temporal requirements are most important for bio-health ontologies. We conduct an empirical investigation of the OBO Foundry using a bespoke methodological approach by searching each of its ontologies for specific *temporal features* and go on to calculate the importance of these features using a sophisticated set of measures. By doing so, we derive a formal set of Temporal Requirements which act as a set of guidelines which a language or logical extension to OWL 2 would need in order to meet the temporal requirements of bio-health ontologies.

Keywords: ontology, bio-health, temporal, OWL, description logics, requirement

1. Introduction

The Web Ontology Language (OWL), as standardised by the World Wide Web Consortium (W3C) is a collection of knowledge representation languages designed for use in many application scenarios, providing the means to model information in a precise and structured way to enable the semantic web. An OWL Ontology is a set of axioms describing the classes and properties of a domain of interest. OWL 2 [1] is the current iteration (and successor) of OWL, and has two levels of expressivity: OWL 2 DL and OWL 2 Full, the former having a Description Logic (DL) as its logical basis. DLs [2–4] are decidable fragments of First Order Logic and have the ability to reason with information in a meaningful way. Two of the main aspects of DLs are to: (1) provide ways to model relations between three kinds of entities in the domain of interest, those being concept descriptions, roles and individuals names and (2) to build complex terms, usually called concept expressions, axioms and assertions and even knowledge bases (or ontologies). There are many varieties of DLs and they often differ by what constructors, axioms and operators are allowed, which in turn offers different levels of expressivity. The DL underlying OWL 2 DL is $\mathcal{SROIQ}(\mathcal{D})$ [5]. Using DLs as the underlying formalism for OWL ontologies come with many advantages. Due to precise syntax and semantics of DLs, they come with the ability to infer new information without having to state it explicitly. OWL (or DL) Reasoners are computational systems that can compute and infer new information from ontologies. Many reasoning services exist depending on what information needs to be deduced. Although many DLs such as $\mathcal{SROIQ}(\mathcal{D})$ are of high complexity (2ExpTime-Complete []), many optimisations have led to efficient implementations of reasoners that have become usable in practice. Lightweight DLs exist with lower complexity levels such as the DL $\mathcal{EL}$, which has polynomial time complexity whilst remaining expressive enough for many ontology applications [6, 7].

The importance of ontologies has increased over the past decade, particularly with applications within the semantic web and life science domain. If we shift our attention solely on applications within life science, particularly those focussed around the bio-health domain, we see a plethora of current ontologies serving different purposes, ranging from describing the development of biological entities, classification of diseases, anatomy descriptions, life cycle stage sequencing and many more. Take the OBO Foundry [8] as an example, an active ontology corpus which has been developed over the past 10 years, containing over 130 actively maintained biomedical ontologies. The corpus contains ontologies such as the Drosophila Gross Anatomy Ontology [9] which describes the anatomy and development of the common fruit fly, as well as medical terminological systems such as the National Cancer Institute Thesaurus (NCIT) [10].

Many applications in life science often include concepts involving time. Take for example an ontology describing the development of some biological entity. Any development inherently involves time: statements made in the ontology could include descriptions of elements developing, an entity occurring during a particular time or an event occurring before, after or during another event. It is clear that time information would be essential in such examples. From a different viewpoint, for instance, in a clinical setting, other temporal information may be needed such as disease progression or medical frequency. Apparently, different application domains embed various types of *temporal features*.

As expressive as ontologies and their underlying DLs are, there are still limiting factors over what they can and cannot express. OWL 2 does offer a way to encode some temporal information, for example, through time stamping (data types), but offers no way to describe any real type of change since as it is still a *static* logic (being a fragment of First Order Logic). It could be beneficial to have some sense of time encoded into the underlying rationale, allowing us to represent and query knowledge in the past, present or future. Clearly, if temporal information is needed but cannot be represented, then it may be the case that many ontologies may be currently misrepresented, or at least OWL does not have the required expressivity to meet the temporal requirements of these ontologies. Currently, it is not clear exactly what kind of temporal expressivity is necessary to meet the temporal needs of bio-health ontologies, simply because the temporal requirements of these ontologies as a whole, are rather diverse and not precisely described.

Many efforts have been made in an attempt to overcome the general problem. Temporal extensions to DLs have been given a lot of attention in recent years. Many proposals exist, ranging from: combining classical temporal logics such as LTL, CTL or CTL* with DLs such as $\mathcal{EL}$ or $\mathcal{ALC}$ [11, 12] where the result can be seen as a two-dimensional Temporal Description Logic (TDL); adding temporal information by extending DLs with a concrete domain [13] to act as a temporal referencing scheme [14]; or even internalising temporal information by embedding it into standard OWL via means of temporal ontologies, for example, a Fluent Ontology [15], or a dedicated OWL Time Ontology [16] which has recently become a W3C recommendation.

Very few of these temporal extensions have been investigated for a specific application area, and those that have are not transferable to other applications. In recent years, research on two-dimensional TDLs has been focussed solely on complexity results rather than capturing the needs of some temporal domain [11], similarly for DLs extended with concrete domains [17]. We believe this is because both have fascinating complexity results [11, 12, 17]: it is very easy for these logics to enter into the undecidability realm, which is undesirable for DLs and ontologies. It may be the case that some of the proposed extensions may, in fact, be suitable for modelling the temporal requirements of bio-health ontologies, but since the temporal requirements of bio-health ontologies are yet to be discovered, an evaluation of these logics has yet to be accomplished. If the requirements were known, we could evaluate the current proposals, to see which were most suited, and if none were, we could set out to define a new logic based on these requirements in an attempt to solve this problem.

In this paper, we provide a foundation for defining a suitable temporal extension to OWL, in particular, to cover the temporal requirements of bio-health ontologies. We produce an empirically validated set of temporal requirements based on a survey of an up to date and actively maintained corpus of bio-health ontologies: the OBO Foundry ontology repository corpus, alongside one of its popular upper level ontologies - the Relation Ontology []. We characterise the corpus with respect to a rich set of *temporal features* and survey their coverage and impact. We then compile a list of *Temporal Requirement Sets*, based on the weighted temporal features. These TRs can be used as either an evaluation mechanism for existing temporal extensions to test their suitability or as a mechanism to drive the definitions of new temporal extensions.

The contributions of this paper are: (1) a temporal encoding of a much used and popular ontology: the Relation Ontology, acting as a seed to our survey, (2) a generalisable entity importance measuring system and (3) a set of TRs acting as a guideline to temporal extensions to OWL.

2. Temporal Patterns in Bio-Health Ontologies

The background and motivation of this paper are presented via examples of how temporal information is currently represented in bio-health ontologies. To be able to do so, we introduce several key biological notions and terms crucial to understand the presented examples. We also introduce key aspects that are relevant to our survey that go hand in hand with temporal modelling. From this point onwards we assume the reader to be familiar with OWL and have a *basic* understanding of Description Logics (DLs), including their syntax and semantics.

2.1. The OBO Foundry

The OBO Foundry [8], first founded in 2007 contains a corpus of ontologies in the biomedical domain. It originally included only 16 ontologies and is to this day a collaborative experiment to establish a set of standards for ontology development, for which they could be used as reference ontologies in the biomedical domain. The corpus now contains over 130 ontologies. The OBO Foundry is home to popular ontologies that range from describing anatomies and developments of organisms such as the Zebrafish, Xenopus, Cephalopod and Drosophila ontologies to those that describe cellular and molecular structures such as the Cell or Gene ontologies. As well as those ontologies that intend to describe some particular domain area, there are those that intend to act as a shared resource, or a *formal structure*, designed to act as a referencing scheme for the domain ontologies for which they can reuse or derive their terms. These ontologies are often referred to as *upper level* ontologies. Two very popular ontologies that are present in the OBO Foundry which fall under this category are the Basic Formal Ontology [18] and the Relation Ontology [19].

The Basic Formal Ontology The Basic Formal Ontology (BFO) is a formal upper-level ontology based on tested conventions for ontology creation. The ontology is built upon a collection of sub-ontologies: the SNAP ontology and the SPAN ontology. The former defines entities known as *continuants* (or *endurants*) and the latter defines entities known as *occurrents* (or *processes*).

In general, continuants are known to be objects that endure or persist through time. They can undergo changes, inhere in objects, be physical objects themselves, but must persist during the times they exist. Examples of continuants are you, your clothes, a pen, a phone, etc. From a biological viewpoint, continuants could include cells, your heart, your blood, your blood type, etc. BFO divides continuants into three separate categories, namely: independent continuants, generically dependent continuants, and specifically dependent continuants. Independent continuants are those continuants that can stand alone and continue to persist, i.e., they do not rely solely on something else for their existence. Dependent continuants do rely on something else for their existence to persist. The difference between specifically dependent continuants and generically dependent continuants is that the former relies on exactly one independent continuant (its bearer) for its existence (and it will cease to exist once its bearer does), whereas the latter can have multiple bearers. An example of specifically dependent continuant is the shape of a ball (round). An example of a generically dependent continuant is an entry in a database (it relies on each value in the entry).

Occurrents, on the other hand, are disjoint from continuants. Occurrents are those entities that unfold through time in temporal phases. They are often referred to as *events* or *processes*. If a continuant were subject to an event occurring, such as a heart (the continuant) beating (the event), the occurrent would be the event itself. Therefore, occurrents are not physical objects themselves; they are the events that unfold around the objects, subject to time. The occurrent class is also partitioned into several subclasses, namely: *process, process boundary, spatiotemporal region* and *temporal region*. A process is an occurrent that has temporal parts and depends on some material entity for some time. For example, consider a person over the course of his life, starting in childhood and ending in late adulthood. The process experienced by this individual would have been the process of ageing, and it would depend on that person itself. Process boundaries are temporal parts of processes that themselves have no other temporal parts. The example given by BFO of a temporal boundary is *"the boundary between the 2nd and 3rd year of your life"*. Temporal regions are simply occurrents that have references to some notion of time (instances or intervals). Examples include the time right now, the range of time during when you were born until your eventual death, the time that covered the year 1990, etc. Finally, spatiotemporal regions are defined as occurrents that are part of spacetime. Examples are the region occupied by the life of a biological entity and the region occupied by the development of a disease.

It is clear that both continuants and occurrents are objects that require time to be defined and understood. Many of the ontologies in the OBO Foundry have incorporated the BFO's class hierarchies into their structures (adhering to OBO's principles), inheriting their properties and definitions. Having a unified and well-defined structure leads to less ambiguity in their understanding and helps to make integration easier.

The Relation Ontology The Relation Ontology (RO) acts as a means for standardisation across ontologies in the OBO Foundry and the wider OBO library. Its main focus is the classification of relations between instances of classes that exist in the biomedical domain, but more importantly, it covers relations used in OBO Foundry ontologies. First introduced in 2007, the ontology was host to only ten relations, including primitive biological relations such as *part of, derives from* and *preceded by*, where each was equipped with a precise definition to avoid any ambiguity of their correct usage. The current version of RO is now host to 497 relations (as of 5th December 2016), where similar levels of detail are used in the definitions for many of the relations. As well as modelling relations, it also comes equipped with a class hierarchy that intends to classify the domains and ranges of the relationships, most importantly, between continuants and occurrents. Specifically, it aligns these classes with those from BFO. As stated, many of the relations in RO come with definitions to avoid ambiguity in their meanings. Some also come with temporal additions in their definitions. Take for example the definition and additional clarificatory comments provided for the mereotopological relation *part of*:

> *"a core relation that holds between a part and its whole"*

> *"Parthood requires the part and the whole to have compatible classes: only an occurrent can be part of an occurrent; only a process can be part of a process; only a continuant can be part of a continuant; only an independent continuant can be part of an independent continuant; only an immaterial entity can be part of an immaterial entity; only a specifically dependent continuant can be part of a specifically dependent continuant; only a generically dependent continuant can be part of a generically dependent continuant. (This list is not exhaustive.)"*

> *"Occurrents are not subject to change and so parthood between occurrents holds for all the times that the part exists. Many continuants are subject to change, so parthood between continuants will only hold at certain times, but this is difficult to specify in OWL."*

The definitions are explained well enough for the terms not to be taken ambiguously, but more importantly, they give information on how they should be interpreted with respect to time (not only by what we can infer from the respective domain and range types) and also show the lack of temporal support from OWL itself.

RO relations cover the vast majority of pairings over the classes they define. For example, relational hierarchies present in RO cover relationships between independent continuants and processes, outlined in the hierarchy *relation between structure and stage*, which include relations such as *existence starts during* and *existence ends during*. Other branches of the hierarchy include relations between independent continuants and specifically dependent continuants such as the relation *bearer of*.

Both occurrents and continuants are crucial to the relations of RO, and thus to all of the ontologies in the OBO Foundry that use RO. As with the BFO, many terms in RO have temporal information present and require this information to be correctly interpreted.

2.2. Temporal Modelling in the OBO Foundry

We now present an example of temporal modelling present in an OBO Foundry ontology. The example will use relations from RO and entities that correspond to those described in BFO and will illustrate the temporal weakness of OWL and show support for our survey.

The Drosophila Gross Anatomy Ontology describes the anatomy and developmental stages of the life cycle of the Drosophila Melanogaster (the common fruit fly). We present a small fragment of the ontology describing the development of the *spermatid cell*, a part of the *male germline cell* of the fly itself. The fragment shows temporal patterns through two of its most used properties; *develops from* and *part of*, and can be broken down between 4 stages shown in the following axioms:

$$LeafbladeS \sqsubseteq \exists \mathbf{dF}.OnionS$$

$$Onion\mathbf{S} \sqsubseteq \exists \mathbf{dF}.ClewS$$
$$Clew\mathbf{S} \sqsubseteq \exists \mathbf{dF}.Agglomeration\mathbf{S}$$
$$Agglomeration\mathbf{S} \sqsubseteq \exists \mathbf{dF}.Coalescence\mathbf{S}$$
$$Leafblade\mathbf{S} \sqsubseteq \mathbf{S}, Onion\mathbf{S} \sqsubseteq \mathbf{S}, Clew\mathbf{S} \sqsubseteq \mathbf{S}$$
$$Agglomeration\mathbf{S} \sqsubseteq \mathbf{S}, Coalescence\mathbf{S} \sqsubseteq \mathbf{S}$$
$$\mathbf{S} \sqsubseteq \exists \mathbf{partOf}.\mathbf{S}Cyst$$
$$(\mathbf{S} = Spermatid, \mathbf{dF} = developsFrom)$$

The first nine axioms express a Spermatid cell going through 5 stages of development (for now we will assume that this short example encodes the entire developmental pattern and nothing occurs before or after the first and last stage). The tenth and final axiom expresses that every Spermatid is part of a Spermatid Cyst. We choose to interpret the identity of the Spermatid cell as the same cell over each developmental stage. Of course, each cell is a distinct element, representing a changed version of its predecessor continuously developing its morphology over time, but when a Coalescence Spermatid develops from an Agglomeration Spermatid, the Agglomeration Spermatid ceases to exist as an entity. In this example at least, we take develops from to represent a specific type of change, which is also apparent in the definition of *develops from*. Again, specific to this example (and others), the *develops from* relation could also be seen to describe both pre and post-conditions of elements' development. For example, in the first axiom, the class Agglomeration Spermatid could describe the precondition and the class Coalescence Spermatid could describe the postcondition of the same element developing. Finally, since the same Spermatid is continuously changing, then each type of Spermatid should belong to the same Spermatid Cyst during its development.

We identify two major temporal aspects of this development sequence. The first is that there is a single entity developing (the spermatid - a continuant) and the second is that there is a continuous partonomy between the two entities (the other element being the spermatid cyst - also a continuant) whilst they are developing. Due to the way the ontology is modelled, none of these temporal constraints can truly be enforced in OWL. Consider Figure 1. The use of the existential restriction '$\exists$' in the axioms may refer to distinct elements for each possible Spermatid, immediately losing any possible identity constraints. This could lead to problems involving errors in the duplication of properties. For example, the Spermatid could have constraints on it itself, and thus each Spermatid in the example model would also be subject to these constraints. Then, if a change was to occur in one Spermatid, it would not necessarily appear in another Spermatid since they could all be distinct. A knock-on effect is that Spermatid Cysts that the Spermatids are part of do not have to be the same Spermatid Cyst, which can again lead to similar problems. In an ideal setting, the identity between the Spermatids must be maintained, as should the partonomy between the same elements. A more *faithful* model is also presented in Figure 1. In this model, we imagine OWL to have an embedded timeline, where we can view *normal* OWL worlds (or models) at different time points, like the two-dimensional semantics seen in LTL_{DL} combinations such as $\text{LTL}_{\mathcal{ALC}}$ [11, 12]. They are called two dimensional since they extend the standard DL domain (the first dimension) with a timeline (the second dimension), and models can be viewed as sequences of standard DL models, that can share the same domain. We adopt a similar approach as it suits this example well. In this temporal setting, there are 5 OWL worlds that are set along a timeline, and each world shares the same 2 domain elements which represent the Spermatid and the Spermatid Cyst. At each time the Spermatid element belongs to a different Spermatid class, for example, at time t the element is an instance of Agglomeration Spermatid class and at $t+1$ it is an instance of Coalescence Spermatid Class. During each time point, the domain element has a *part of* relation to a Spermatid Cyst, which is the same Spermatid Cyst throughout the development. Such a model seems to capture more faithfully what was intended for the biological modelling, yet this type of modelling is beyond OWL. There is only one single world of evaluation, no timeline, and no identity constraints between distinct entities.

This example shows yet another clear-cut case of OWL's lack of temporal expressivity, and more importantly shows a significant amount of temporal information loss for only two relations and a small number of axioms. The motivation of this paper is driven by examples such as these; *develops from* and *part of* alone seem to be important relations for the Drosophila Ontology. Together, they are roughly used in one-third of the total logical axioms in the ontology, which could imply that one-third of the ontology is unfaithfully modelled. It would also be useful to know how often they are used in other bio-health ontologies. If they are only used in the Drosophila Ontology and no other, then it would be an over statement to say that both of

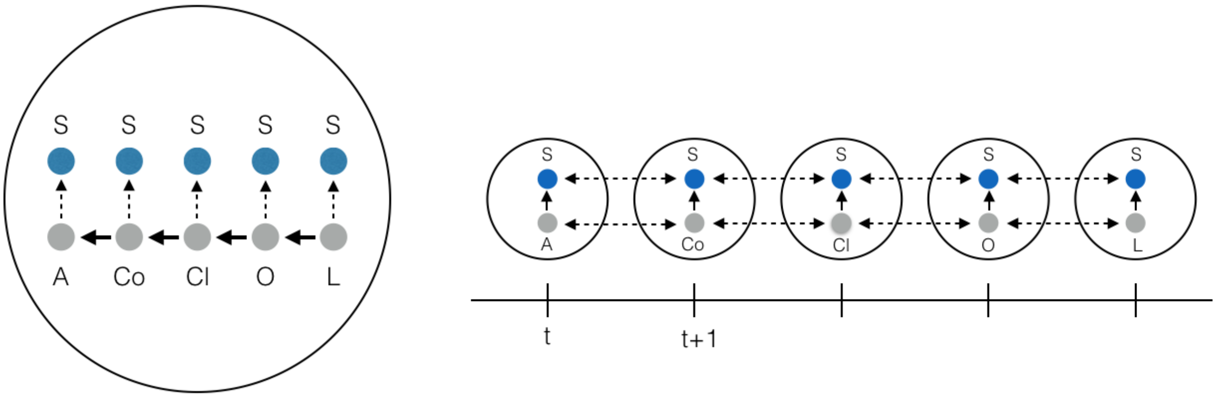

Fig. 1. Left: An OWL model of a development fragment of the Drosophila ontology. Right: A temporalised OWL model of the same development fragment. ○=element of the DL domain, ←=develops from, ↑=part of, ←−−→=identity relation, S=Spermatid Cyst, A=Agglomeration, Co=Coalescence, Cl=Clew, O=Onion and L=Leafblade, X=class name

the relations were of crucial importance to the temporal modelling of bio-health ontologies. Yet, if they were also used in one-third of axioms in all bio-health ontologies, it would not be unfair to say they were important relations. It would also not be unfair to state that, for example, independent continuants were important for modelling in bio-health ontologies, since the domain and range of *develops from* are restricted to this specific class, which would mean that one-third of the axioms in those ontologies require independent continuants.

The relations *develops from* and *part of* encode specific temporal information: *develops from* relates entities over two different time points (a past time relation), whereas *part of* relates entities in a single time point (a same time relation). Moreover, *develops from* relates two independent continuants, where as *part of* can be used for continuants or occurrents, provided both types are compatible. We call these attributes of relations *temporal attributes*. Using the same reasoning as above, all of these attributes could be seen as important for temporal modelling of bio-health ontologies. If there was another relation in the Drosophila ontology that had the same temporal attributes as *develops from* that was also considered important, then it would make sense to also focus on the importance of the attributes themselves rather than just the individual relations.

Our survey intends to empirically and systematically rank the importance of these types of temporal features. We propose to annotate all relations in RO that are used across The OBO Foundry with their temporal attributes and then use carefully designed metrics to define their importance using their logical axiom counts and more. Such analysis will give rise to a set of temporal requirements of those bio-health ontologies.

We now go on to explain how the temporal attributes are derived and present the definitions of the metrics used to define importance.

3. Materials & Methods

In the following, we distinguish three types of temporal features: (1) Temporal relations are those (RO) relationships that encode information that is temporally relevant; (2) Temporal attributes are types of temporal information that represent temporal phenomena described by temporal relations, and (3) Temporal annotations are sets of temporal attributes. (2) and (3) are defined in detail in the following section.

A temporal requirement corresponds to a temporal annotation. For example, if annotation A is used in an axiom of an ontology, A is said to be a temporal requirement of that ontology. Lastly, a temporal requirement set is a set of temporal requirements, typically one where the temporal requirements are likely to co-occur, defined in more detail in the following.

3.1. Overview

The goal of our study of temporal requirements of bio-health ontologies is two-fold. First, we will study the importance of temporal features across OBO Foundry ontologies. Second, we will suggest an empirically validated, ordered list of temporal requirement sets. In order to achieve our goal, we:

1. Define a set of temporal attributes based on relations from the RO that are used across the OBO Foundry.
2. Match axioms across the OBO Foundry ontologies which exhibit these attributes using a smart matching technique.
3. Analyse the resulting data with respect to the importance of these attributes and their corresponding temporal annotations.
4. Derive a ranked list of temporal requirements based on the importance, coverage and necessity score of temporal annotations across the OBO Foundry corpus.

3.2. Defining and Identifying Temporal Attributes

We use the relationships defined in the relation ontology (RO) as a source for defining and extracting temporal attributes. We define temporal attributes as types of temporal information that represent temporal phenomena described by RO relations, such as the *past time relation* phenomena found in the *develops from* relation. For each relationship, the temporal information is gathered from its definitions or other annotations, its domain and range constraints, related relationships due to OWL's precise semantics and in some circumstances general biological knowledge and the way in which ontologies use the relationship when the first three may be lacking.

To illustrate this procedure, recall the RO relationship *part of.* As well as the annotations (including definitions) presented in Section 2.1, take as well the annotation

"axiom holds for all times"

meant to be temporally interpreted as *"if x:C, and C has part D, then for all times t, x will always have a part y:D"*. The temporal information we gather from the *part of* relationship is that (1) partonomy relationships take place during single time points, i.e. they are same-time relations, (2) the classes must be compatible, (3) partonomy may hold eternally true (when the elements exist) and (4) the partonomy may hold between the same elements over time. (4) is also derived from the fact that in many temporal modelling scenarios, it may be important that the same elements are related over time. For example, if a particular cell were to have a nucleus as a part at some time point, it would not make sense for this cell to have another nucleus at another time point in usual cell development patterns (this is often referred to as a rigid relation in the temporal logic realm). These temporal features are then categorised into the following respective temporal attributes (1) ***Time:same*** indicating the relation takes place over a single time point, (2) ***Domain:X-Range:X*** indicating the domain and range must share the same type *X* (where *X* is either a type of continuant or a type of occurrent), (3) ***AHFAT*** (**A**xiom **H**olds **F**or **A**ll **T**imes) and (4) ***Rigid*** indicating the relations follow a rigid like pattern.

We performed this temporal attribute derivation procedure for every RO relationship used amongst ontologies in the OBO Foundry. We acquired 56 distinct temporal attributes which we categorised into the following 6 sets: (1) ***Domain & Range***, (2) ***Time***, (3) ***States***, (4) ***Identity***, (5) ***Rigid***, (6) ***AHFAT***.

Domain & Range contains the set of all pairings of domain and range constraints that occurred in RO relationships. The set contains 23 attributes involving the four types of continuants, general occurrents and processes. Eight of the attributes are between continuants and occurrents (e.g., *Domain:C-Range:O* or *Domain:O-Range:C*), 11 are between only continuants (e.g., *Domain:IC-Range:IC*), two are between only occurrents (e.g., *Domain:O-Range:O*), one was between any element and a continuant (e.g., *Domain:X-Range:C*) and one was between any two elements of the same kind (e.g., *Domain:X-Range:X*).

Time contains attributes describing how each relationship relates its entities in time. Due to the fundamental temporal differences between continuants and occurrents, the set can be partitioned into three subsets, those being time attributes of relations between two continuants, two occurrents, or between continuants and occurrents. Overall this set consists of 19 attributes. The continuant time attributes account for seven of these, consisting of *Time:same, Time:diff, Time:past, Time:pastImmediate, Time:same/past, Time:future* and *Time:same/future*. *Time:same* indicates that the domain element of a relationship is related to the range element at the *same* moment in time. *Time:past* indicates that the domain element of a relationship is related to the range element present at a *past* moment in time. *Time:pastImmediate* indicates that the domain element of a relationship is related to the range present at the *previous* moment in time. *Time:same/past* indicates that the domain element of a relationship is related to the range element present at either a *previous* moment in time or the same moment in time and so on. *Time:diff* is the opposite of *Time:same*, indicating that the domain and range element are in different time points. The occurrent time attributes adopt Allen's time relations on intervals [20]. 13 attributes make up this sub group consisting of *Time:before, Time:before/during, Time:beforeInverse, Time:during, Time:during/overlaps, Time:during/overlapsInverse, Time:finishes, Time:finishesInverse, Time:isEqualTo, Time:meets* and *Time:meetsInverse*. *Time:before* indicates that the domain element of the relationship happens entirely before the range element, where the *before* is to be interpreted as Allen's interval relations intends, i.e., the domain ends before the range starts. *Time:during/overlaps* indicates that the domain element either happens during the range element or overlaps the range element, and so on. Relations between continuants and occurrents are simply a subset of those between continuants. The set consists of the following four attributes: *Time:same, Time:same/future, Time:future* and *Time:same/past*, interpreted in the obvious way.

States contain attributes that describe possible state changes of either the domain or range element of the relationship. Six attributes are contained within this category consisting of *Domain:Birth, Domain:Changed, Domain:Death, Range:Birth, Range:Changed, Range:Death* and *Range:Death*. *Domain:Birth* indicates that the relationship specifies the start of the domain element's existence. *Domain:Changed* indicates that the domain element goes through some type of change (such as a change in class or other properties) compared to what it was previously. *Domain:Death* indicates that the relationship specifies the end of the domain elements existence. The same holds for the *Range:X* attributes in relation to the range elements.

Identity consists of only a single attribute *Identity:same* which indicates that both the domain and range element of the relationship share the same identity, i.e., they represent the same temporal entity.

Rigid consists of only a single attribute *Rigid* which indicates that the relationship follows one of a rigid pattern, where both the domain and range elements of the relationship are required to be consecutively related through time for some required duration.

AHFAT consists of only a single attribute *AHFAT* which indicates that the relationship's domain element is required to have a relation to a compatible range element at all times (during its existence).

Each attribute may also be paired with a tag *Necessary:No* which indicates that it is not necessary for the corresponding relationship to hold that particular attribute, although in some scenarios it can. For example, the attribute *Rigid-Necessary:No* is interpreted as *"it is not necessary in all cases for the relation R to be interpreted rigidly, but in some cases, a rigid interpretation holds for R"*. An example of where this may be the case is where an ontology specifically describes a temporal information.

Hierarchical relationships exist between many of the temporal attributes, since some of the attributes imply others. Figure 2 shows how each attribute type is positioned in its corresponding hierarchy. *The Domain & Range* attributes are ordered depending on their ontological constraints according to the RO class hierarchy and the time attributes are ordered according to their temporal relation types.

3.3. Temporal Annotations

With the resulting temporal attributes, we developed a coding scheme to then annotate each RO relationship with what we call a *temporal annotation* which consists of its temporal attributes, defined as follows:

Definition 1 (Temporal Annotation)**.** *Let R be a relation from RO, and Y ={**Domain & Range, Time, States, Identity, Rigid, AHFAT**} be the sets of temporal attributes described above. A temporal annotation for R is a set $A \subset \cup Y$ where A contains*

- *a single domain and range attribute*
- *0 or 1 identity attributes*
- *a single time attribute*
- *0 or 1 rigidity attribute*
- *1 or more state attributes*
- *0 or 1 AHFAT attributes*

For the purpose of exhausting the temporal attribute completeness in each RO relationship, we also annotate each RO relationship with its implied attributes according to the temporal attribute hierarchy in Figure 2, with what we call a *temporal inferred annotation*, defined as follows:

Definition 2. *Let R be a relation from RO with an existing temporal annotation A. Let $(P, \leqslant)$ be the poset shown in Figure 2. The* temporal inferred annotation

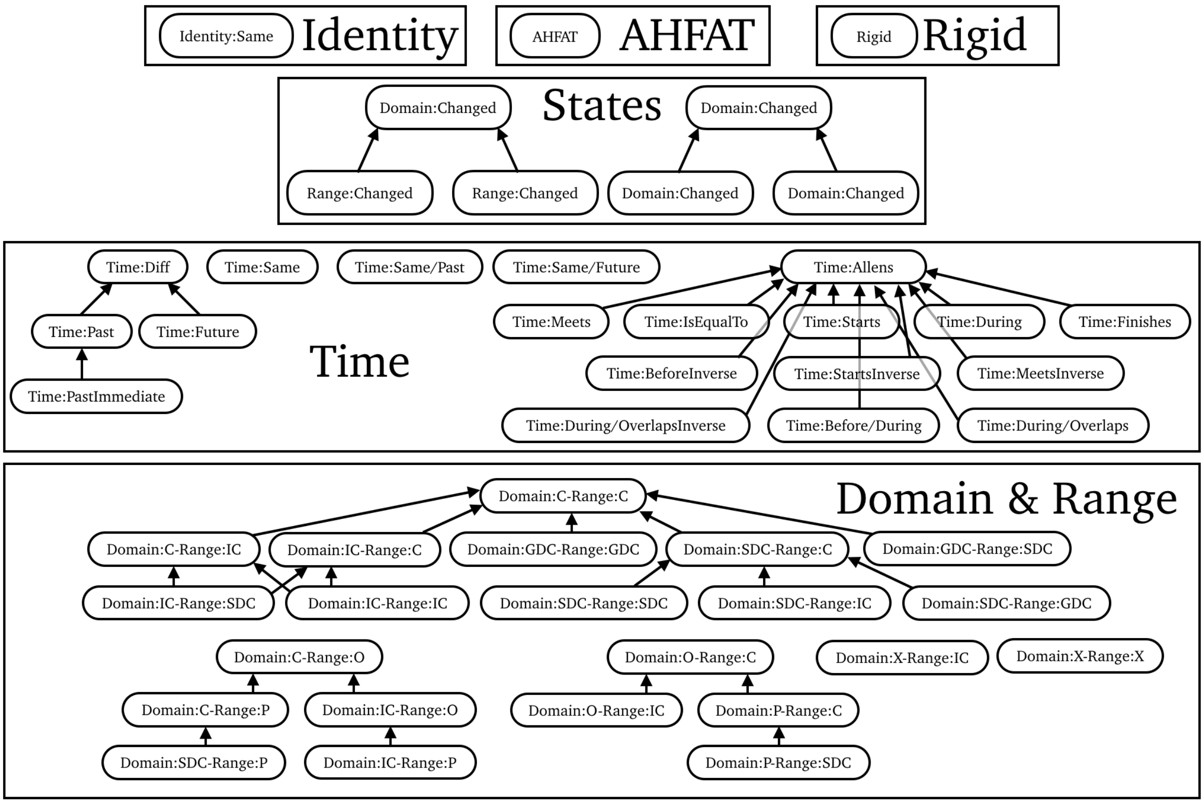

Fig. 2. Hierarchy of temporal attributes.

for R, represented as the closure cl of A, is defined as follows:

$$cl(A) = \{y \mid \exists x : x \in A \wedge x \leqslant y\}$$

The *Necessary:No* tags do not necessarily have to appear on the inferred attributes. As an example, the temporal annotation A_1 for *part of* is {*Domain:X-Range:X, Time:same, AHFAT, Rigid-Necessary:No*}. Its temporal inferred annotation A_1^I is equal to A_1. The temporal annotation A_2 for *develops from* is {*Domain:IC-Range:IC, Time:past, Identity:Same-Necessary:No, Domain:Birth-Necessary:No, Domain:Changed*}. Its temporal inferred annotation $A_2^I = A_2 \cup$ {*Domain:IC-Range:C, Domain:C-Range:IC, Domain:C-Range:C, Time:diff*}.

3.4. Matching temporal features across OBO foundry ontologies

Although the rules of the OBO Foundry enforce that terms, such as relationships, be used consistently throughout (at least) OBO Foundry ontologies, there are instances where this is not the case. Ideally, to check for a relationship's usage in an ontology, one should be able to simply search the ontology's signature for an occurrence of the relationship's IRI. However, this relies heavily on ontology developers *correctly* using terms from other vocabularies, i.e. importing vocabularies. This is often not the case since importing ontologies could result in negative side effects such as size increase or a jump in complexity. In the RO case, this matter is immediately realised. Its expressivity is very high due to its complex modelling of relations (role hierarchies, role chains, size, etc) and importing the RO will most likely have a direct negative effect on performance and reasoning time. If not importing the ontology, then at the least the same IRI of any relation used should be adopted in order to indicate the intention that the relationship is the same relationship from RO. Unfortunately, this is not always the case. Instead, developers may (and do) create their own entity with a similar name. For this reason, we cannot simply rely on checking for exactly matching IRIs in an ontology's signature. Therefore, we adopt a *smart matching* approach, where we define that a relationship outside RO *smartly matches* a RO relation if either they share the same IRI, name (rdfs:label), alternative term (IAO_0000118), OBO foundry unique label (IAO_0000589) or the same exact synonym (hasExactSynonym) to avoid any potential misses. These annotation properties were chosen due to the information encoded in each: they are clear, unambiguous in their meaning and ontologies that define their own relationship would be likely to use values from these annotations. Manual inspection of the annotation properties' values and self-defined relations in the RO confirm this. *Exact matches* occur when a relationship inside an OBO ontology has the same IRI of a relation from RO (i.e., exact matches refer to the correct usage of RO relations in external ontologies).

3.5. Usage of temporal features

We present a notion of *usage* that defines if and how an ontology in OBO *uses* a temporal attribute, annotation or relationship from the relation ontology.

When considering usage throughout the corpus, we shift our attention towards the terminological aspects of the ontologies in the corpus. That is, we choose to investigate the explicitly asserted terminological knowledge, specifically TBox axioms. Our notion of usage is defined as follows:

Definition 3. *Let f be a temporal attribute, F a temporal annotation, P an RO relationship, $\mathcal{O}$ an ontology occurring in the OBO Foundry and let α be a terminological axiom in $\mathcal{O}$. We say that*

- *F **uses** f if $f \in F$*
- *P **uses** F if P is annotated with F*
- *α **uses** P if P occurs in α*
- *$\mathcal{O}$ **uses** P if P occurs in $\mathcal{O}$*

*where **uses** is transitive.*

3.6. Analysing the importance of temporal features

Our goal is to determine the importance of temporal attributes, relations and annotations. To date, no agreed-upon measure exists to quantify the importance of a particular entity $\mathcal{E}$, such as a relation or a class, neither in the context of a single ontology nor across an entire corpus. Entities in an ontology can be used in a variety of ways: they can be used to define the logical content of an ontology, for example in the definition of classes or other logical axioms, or even non-logical expressions such as annotations. As we are interested in determining the requirements for temporal extensions to a knowledge representation formalism, we care only about how entities are used across logical axioms (Def. 3).

To quantify the importance of a particular temporal feature, we decided to rely on *coverage* and axiom usage, which we refer to as *impact* for brevity. We define both metrics for temporal features as follows:

Definition 4. *Let e be either an attribute or annotation and $\mathcal{C}$ be a set of ontologies.*

$$Coverage(e) = \frac{\#\{\mathcal{O} \in \mathcal{C} \mid \mathcal{O} \text{ uses } e\}}{\#\{\mathcal{O} \in \mathcal{C}\}}$$

$$Impact(e) = \frac{\sum_{O \in \mathcal{C}} \left(\frac{\#\{\alpha \in \mathcal{O} \mid \alpha \text{ uses } e\}}{\#\{\alpha \in \mathcal{O}\}} \right)}{\#\{\mathcal{O} \in \mathcal{C}\}}$$

The coverage measures how many ontologies each feature is used in at least once. The impact describes the percentage of axioms a feature occurs in per ontology (note that we present both metrics as proportions over the whole corpus). As previously discussed, neither measure can perfectly quantify importance alone, therefore, we use both in our analysis where appropriate. In our survey, we will determine the impact and coverage of all temporal relations identified through smart matching, as well as the impact and coverage of their temporal features across the OBO Foundry ontologies. We also define a score to quantify the overall *importance* of a feature, which takes into account both the coverage and the impact, defined as follows:

Definition 5. *Let e be a temporal feature and $\mathcal{C}$ be a set of ontologies.*

$$Importance(e) = \frac{n(Coverage(e)) + n(Impact(e))}{2}$$

where n() is a normalisation function that linearly rescales the data values to a range between 0 and 1.

The normalisation $n()$ is applied to give both coverage and impact equal weight towards the importance score.

3.7. Ranked list of temporal requirement sets

Our goal is to produce an ordered list of temporal language requirements based on the results of our survey. We define a temporal requirement set, denoted $\mathcal{R}$, as a set of temporal annotations. For example, the temporal knowledge in $\mathcal{O}$ requires $\mathcal{R}$ if $\mathcal{O}$ *uses* every annotation $\mathcal{A}$ in $\mathcal{R}$. In order to quantify the *Importance* of $\mathcal{R}$, we make use of the following three metrics: *Coverage*, *Necessity* and *Mean-Annotation-Importance*. *Coverage* corresponds to the number of ontologies that can be fully expressed if the temporal requirements in $\mathcal{R}$ are met (i.e., the set of all temporal annotations used in $\mathcal{O}$ is a subset of $\mathcal{R}$:

$$Cov(\mathcal{R}) = |\{\mathcal{O} \in \mathcal{C} \mid \forall \mathcal{A} : \mathcal{O} \textit{ uses } \mathcal{A} \textit{ implies } \mathcal{A} \in \mathcal{R}\}|$$

This metric is of particular interest to language developers whose goal is to enable as many knowledge engineers as possible to express the full set of their temporal requirements. The disadvantage is that covering requirement sets are often large, i.e. contain a large number of temporal annotations and attributes, and may, therefore, be difficult to realise.

(2) The *necessity score* corresponds to the number of ontologies that need a particular set of temporal requirements to be met, i.e. $\mathcal{R}$ is a subset of the set of all temporal annotations $\mathcal{A}$ used in $\mathcal{O}$:

$$Nec(\mathcal{R}) = |\{\mathcal{O} \in \mathcal{C} \mid \forall \mathcal{A} \in \mathcal{R} : \mathcal{O} \textit{ uses } \mathcal{A}\}|$$

The advantage of using this metric as the basis for language design is that requirements with a high necessity score are typically small, and may benefit a wider group of users. The disadvantage is that there is no guarantee that any user will have all of their temporal requirements satisfied (or indeed a significant proportion).

(3) The third metric, *mean annotation importance*, is the mean importance score (see Definition 5) of all annotations in the requirement set:

$$MeanAnnImp(R) = \frac{\sum_{\mathcal{A} \in \mathcal{R}} Importance(\mathcal{A})}{|\ \mathcal{R}\ |}$$

To quantify the overall importance of a requirement set, we use the following formula:

$$Importance(\mathcal{R}) = \frac{n(Cov(\mathcal{R})) + n(Nec(\mathcal{R})) + n(MeanAnnImp(\mathcal{R}))}{3}$$

The normalisation function $n()$ is used for the same reason as in Definition 5. As the total requirements space is in the worst case exponential in the number of distinct annotations[1], we decided to consider only those combinations of temporal annotations that co-occur in some OBO Foundry ontology. For example, if the full set of annotations used in an ontology $\mathcal{O}_1$ was A_1, A_2 and A_3, we considered the requirement $\mathcal{R}_1 = \{A_1, A_2, A_3\}$ for our analysis. This reduces the space of possible requirements drastically (to, in the worst case, the number of OBO Foundry ontologies). The advantage is that we do not have to concern ourselves with combinations of annotations that might be practically useless (because of annotations that would never co-occur in real ontologies). On the flip-side, the converse is true: we might miss small, *almost* covering requirement sets that could be potentially very useful. We do believe however that it is, when in doubt, best to be guided by the empirical distribution of co-occurring temporal annotations, so we chose to restrict our attention to "used" annotation combinations. Following this procedure resulted in a total of 75 requirements.

4. Results

A full account of the analysis (scripts and all results) can be found on **rpubs** (http://rpubs.com/matentzn/obo-tdl-v2). Although our main focus is on determining the importance of temporal requirements, we first discuss the findings of matchings, relations and attributes.

4.1. Smart & Exact Matching

For each ontology, we iterated through each terminological axiom and recorded whether or not the axiom contained an an exact match, or otherwise a smart match of an RO relation. We repeated this for every axiom in every ontology, for every relation in RO.

Out of 140 downloadable ontologies (December 2016) of the OBO Foundry Repository, 11 were not parseable. While 31 ontologies contained no RO relations according to our matching approach, 98 ontologies contained smart matches. It is noteworthy that, if we had relied on exact matches alone, only 68 ontologies would have matched RO relations. This means that we would have underestimated the need for temporal modelling significantly (30% of the OBO Foundry ontologies would have been ignored).

[1]The powerset of all possible annotations

In terms of the axioms the relations are used in, if we were to ignore axioms that only had smart matches, we would be ignoring, again, 30% of all axioms in the OBO Foundry. Of course, it could be the case that all of the smart matches were incorrect matches (they were not meant to simulate RO relations), but we did investigate a reasonably sized random selection of the matches, and it seemed obvious that the relations were matched correctly. For example, some of the matched relations investigated were used in the same way (even temporally) as the way they are defined in the RO. Table 1 shows, for the top 10 elements, by how much the coverage would be underestimated when considering only exact matches.

Relation	Exact	Smart	% Diff
part of	52.04	79.59	52.94
has part	40.82	48.98	19.99
inheres in	24.49	29.59	20.82
has participant	17.35	27.55	58.79
has role	16.33	26.53	62.46
realizes	21.43	24.49	14.28
located in	18.37	21.43	16.66
has quality	12.24	20.41	66.75
bearer of	15.31	19.39	26.65
develops from	16.33	19.39	18.74

Table 1

The top 10 RO relations showing their smart matching and corresponding exact matching metrics in terms of the percentage of ontologies they were matched in. % Diff is the percentage difference between the exact and smart matches.

4.2. Importance of Temporal Features

The temporal features are categorised based on their domain and range type, and analyses are performed within these categories. This decision was made because each feature contains different combinations of temporal attributes, which cannot be meaningfully evaluated against attributes contained in features with different domain and range types. This way, the analyses are rendered more comprehensible, and comparisons may be drawn against similar temporal phenomena. The domain-range categories used are Continuant-Continuant (CC), Occurrent-Occurrent (OO), Occurrent-Continuant and Continuant-Occurrent (OC-CO) and Other (OT) that includes features that contain the attribute (*Domain:X-Range:X*).

4.2.1. Temporal Relations

We begin by providing a short analysis of temporal relations used across OBO Foundry ontologies. The full tables that display the impact and coverage for every matched relation can be seen in Appendix A. A total of 145 relations were used across the OBO Foundry, of which 98 were CC (68%), 24 were OC-CO (17%), 18 were OO (12%) and 5 were OT (3%).

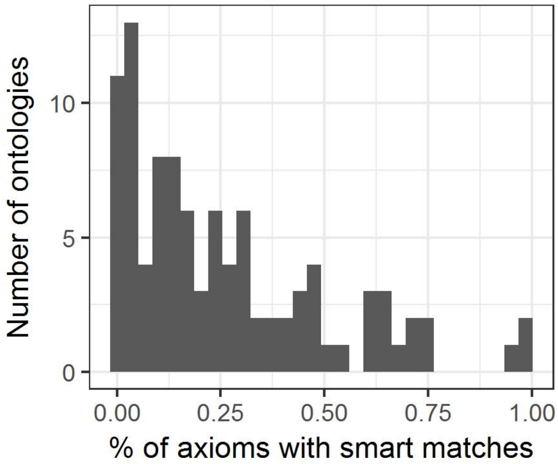

Fig. 3. Distribution of the proportion of axioms with smart matches across ontologies

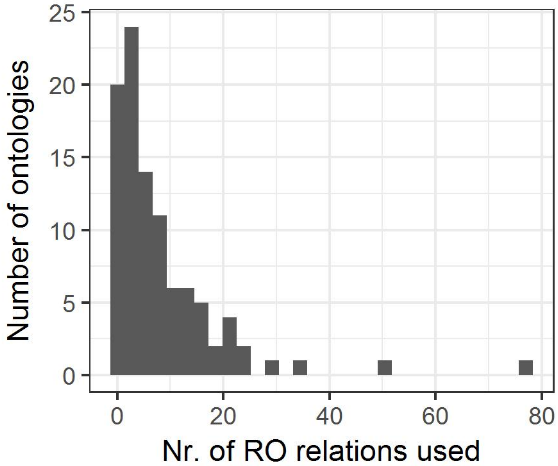

Fig. 4. Distribution of RO relation usage across ontologies

Figures 3 and 4 show two histograms illustrating the prevalence and diversity of relations used. Figure 3 shows the distribution of ontologies by smart match prevalence, i.e the proportion of axioms that use at least one RO or RO-like relation compared to the total number of axioms in the ontology. For example, the microRNA ontology (MIRNAO) has 764 axioms, with 79 axioms using at least one of RO(-like) relation, resulting in a proportion of $79/764 = 10.34\%$. As can be seen, there are 2 ontologies that have near 100% relation usage in their axioms. Most have relation prevalence in the range of $0\% - 75\%$, gradually declining towards the high proportion end. There is a large peak around the 0% region. Some ontologies responsible for this peak are those that have large axioms counts, but low RO relation usage.

Figure 4 illustrates the diversity of RO relations as the total number of different RO relations that were used in an ontology. For example, MIRNAO makes use of 8 different RO relations (which is close to the empirical mean of 8.3 different relations per ontology). Only 8 ontologies contain more than 20 different RO relations, and, perhaps apart from UBERON (78) and OVAE (51), even these contain only a fraction of all existing RO relations. This indicates an overall low diversity of RO relations across single ontologies, however, we believe this to be expected: for an ontology to have a high diversity of relations, the domain for which the ontology covers would be considerably large. The majority of ontologies in the OBO Foundry cover specific areas of interest, ignoring the few upper-level ontologies that intend to classify general knowledge. This can explain both the high coverage across the corpus and the comparatively low within-ontology relation diversity.

Relation	#O	Coverage	CAT
part of	78	79.59	OT
has part	48	48.98	OT
inheres in	29	29.59	CC
has participant	27	27.55	OC-CO
has role	26	26.53	CC
realizes	24	24.49	OC-CO
located in	21	21.43	CC
has quality	20	20.41	CC
bearer of	19	19.39	CC
develops from	19	19.39	CC

Table 2
Top 10 temporal relations ordered by coverage

Relation	Impact	CAT
part of	11.52	OT
has part	3.03	OT
immediately preceded by	2.24	OO
inheres in	2.07	CC
has quality	1.52	CC
bearer of	1.30	CC
develops from	0.99	CC
has modifier	0.65	CC
derives from	0.57	CC
preceded by	0.56	OO

Table 3
Top 10 temporal relations ordered by impact

Summary metrics of impact and coverage can be seen in Table 4. Tables 2 and 3 show the top ten relations amongst all categories, ordered by their coverage and impact respectively. As can be seen in Tables 2 and 3, two OT relations have the highest impact and coverage. The remaining top ten relations for coverage and impact are mostly CC relations, with only 3 relations being OC-CO or OO.

As can be seen in Table 4, the average coverage and impact for CC, OO and OC-CO relations are roughly the same, whereas they are considerably higher for OT. The OT category dominates the relation results. This is due to the relation *partOf* which has both the highest scores by a considerable margin for impact and coverage out of all relations. Its inverse, *hasPart* also contributes to the high scores of the OT category with relatively high scores, outscoring every relation from any other category. The remaining relations in OT have low scores. Although the CC category has the highest number of used relations (98), only 12 have a coverage above 10 with the remaining relations' coverage gradually declining towards 1.02 (1 ontology). Only 3 CC relations have impact above 1. OO and OC-CO have similar trends: few relations have relatively high coverage scores with the remaining declining steadily towards 0, and even fewer have notable impact scores. There is an overall strong correlation between coverage and impact for the CC, OC-CO and OT categories each falling above 0.7, whereas the OO correlation was only 0.552.

Type (n)	μ-cov (σ)	μ-imp (σ)	Correl	min-cov	max-cov	min-imp	max-imp
CC (98)	4.414 (5.802)	0.106 (0.314)	0.764	1.02	29.59	0	2.066
OO (18)	5.555 (5.482)	0.244 (0.532)	0.552	1.02	18.37	0	2.24
OC-CO (24)	6.845 (7.538)	0.127 (0.155)	0.763	1.02	27.55	0	0.56
OT (5)	26.734 (35.949)	2.969 (4.946)	0.94	1.02	79.59	0.005	11.519

Table 4

Metrics of relations ($n = 145$) in each domain and range category

4.2.2. *Temporal Attributes*

Attribute	#O	Coverage	CAT
OT-Dom:X-Ran:X	84	85.71	OT
OT-Rig:Yes-Nec:No	84	85.71	OT
OT-TI:AHFAT	84	85.71	OT
OT-Time:Same	84	85.71	OT
CC-Dom:C-Ran:C	68	69.39	CC
CC-Dom:IC-Ran:C	62	63.27	CC
CC-Time:Same	60	61.22	CC
CC-Rig:Yes-Nec:No	59	60.20	CC
CC-TI:AHFAT	53	54.08	CC
CC-Dom:C-Ran:IC	46	46.94	CC

Table 5

Top 10 temporal attributes by coverage

Attribute	Impact	CAT
OT-Time:Same	14.85	OT
OT-Dom:X-Ran:X	14.55	OT
OT-Rig:Yes-Nec:No	14.55	OT
OT-TI:AHFAT	14.55	OT
CC-Dom:C-Ran:C	10.40	CC
CC-Time:Same	8.49	CC
CC-Rig:Yes-Nec:No	8.22	CC
CC-Dom:IC-Ran:C	6.72	CC
CC-TI:AHFAT	4.83	CC
OO-Dom:O-Ran:O	4.39	OO

Table 6

Top 10 temporal attributes by impact

Coverage & Impact The coverage and impact of all temporal attributes can be found in Appendix B. The top ten attributes for both coverage and impact can be seen in Tables 5 and 6 respectively. OT attributes followed by CC attributes dominate the top ten scores, with only one other attribute from the OO category appearing in the top ten for either metric. The average coverages and impacts for each category have more variation than in the relation case.

73 attributes were used across all domain and range categories with 31(42%) belonging to CC, 16(22%) to OO, 21(29%) to OC-CO and 5(7%) to OT. The correlation between coverage and impact for each category is high ($\mu = 0.898$).

When considering CC attributes, it is clear that the most popular domain and range combinations were those between ICs (domain) and Cs (range). Other combinations are also prominent involving SDCs, whereas relations involving GDCs are less frequent. The *Time:Same* attribute, which indicates that elements involved in the relation are related at the same time point, has both higher coverage and impact than the *Time:Diff* attribute, which indicates that the elements are related at different time points (e.g., *developsFrom*). There is a considerable difference between the two (and for each of *Time:diff*'s subtypes), although the coverage of *Time:diff* is not low enough to ignore. Attributes from the ***States*** set are less frequent, with notable coverages, but low impacts. Finally, the attribute *Rig:Yes* scores in the top 3 attributes for coverage and impact, indicating that the majority of used CC relations require this feature.

OO relations only differ by their ***Time*** and ***Domain& Range*** attributes. Only 4 ***Time*** attributes have coverage above 10, and only one of which, *Time:MeetsInverse* has an impact score above 1. The overall impact average was particularly low for OO attributes. OO relations that were specifically declared to be between processes (identified by those relations having the attribute *Dom:P-Ran:P*) have a coverage of 10.20, roughly 25% of overall OO attribute coverage, but their impact is significantly lower at only 0.157, around 3% of the total impact for OO attributes.

Only 5 OC-CO attributes have impact over 1, with 3 coming from the ***Domain & Range*** set, 1 from the ***Rigid*** set and 1 from the ***State*** set. These attributes also appear in the top scoring coverage attributes. There is no significant *Domain & Range* type attribute that stands out above others. Two noteworthy findings are that (1) the *Time:Same* attribute has both higher coverage and impact than *Time:Diff*, and (2) the *Rig:Yes* attribute plays a key role.

The majority of OT attributes have the highest scores amongst all attributes, which are those that are

Type (n)	μ-cov (σ)	μ-imp (σ)	Correl	min-cov	max-cov	min-imp	max-imp
CC (31)	25.313 (21.226)	2.094 (2.831)	0.921	1.02	69.39	0.001	10.398
OO (16)	11.544 (13.336)	0.834 (1.495)	0.909	1.02	41.84	0	4.393
OC-CO (21)	18.901 (14.032)	0.665 (0.629)	0.761	3.06	46.94	0.066	2.376
OT (5)	69.18 (36.962)	11.76 (6.41)	1	3.06	85.71	0.296	14.849

Table 7

Metrics of attributes ($n = 73$) in each domain and range category

contained within the annotations for the *hasPart* and *partOf* relations. Interestingly, the attribute *Rig:Yes* is one of the most used attributes, in terms of coverage and impact.

4.3. Temporal Annotations and Temporal Requirements

The coverage and impact scores of all annotations can be seen in Appendix C, with summary metrics in Table 8. A list of all annotations can be seen in Table 16 (Appendix C). Tables 10 and 11 show the top ten annotations amongst all categories, ordered by their coverage and impact respectively.

The coverage of annotations in each category follows a similar trend: a fraction of the annotations have coverage above 10, with the remainder gradually declining towards the minimum (1.02). Very few annotations have notable impact scores in each category, only 6 annotations have impact over 1 in the CC, OO and OT categories, and none have impact over 1 in OC-CO.

4.3.1. Analysis of temporal requirements

Requirement sets are sets of temporal annotations. As described in Section 3.7, we restrict our attention to those requirement sets that co-occur in at least one ontology. To quantify the importance of requirement sets, we take a two step approach. First, we compute an overall importance score, introduced in Section 3.7. Second, we compute the Pareto frontier. The reason for this is that the importance score does not account for the fact that, based on the three main metrics: coverage; necessity; and mean annotation importance, the ranking between the requirement sets is a partial order. All requirements sets and their importance scores can be seen in Appendix D, in Tables D.1 and D.2.

Annotation	#O	Coverage	CAT
A68	84	85.71	OT
A32	34	34.69	CC
A38	34	34.69	CC
A63	29	29.59	CC
A57	27	27.55	OC-CO
A59	24	24.49	OC-CO
A43	21	21.43	OO
A2	19	19.39	CC
A26	19	19.39	CC
A39	19	19.39	CC

Table 10

Top 10 temporal annotations by coverage

Annotation	Impact	CAT
A68	14.55	OT
A51	2.24	OO
A38	2.23	CC
A63	2.19	CC
A39	1.30	CC
A26	1.04	CC
A32	0.81	CC
A23	0.76	CC
A65	0.65	CC
A43	0.63	OO

Table 11

Top 10 temporal annotations by impact

75 temporal requirements were identified, of which the top 15 (according to their importance score) can be seen in Table 9. Requirements on the Pareto frontier (14 in total), are shaded in grey (they do not have any requirement sets that are strictly better than them). For example, R10 is not on the Pareto frontier, but ranks third according to our importance score. This is because it scores, taking into account all three metrics, strictly worse than R19, while the overall importance score is still very high. The reason for the requirement sets on the Pareto frontier being scattered across the whole table (see Appendix D for the full version) is that there is naturally a strong trade-off between the

Type (n)	μ-cov (σ)	μ-imp (σ)	Correl	min-cov	max-cov	min-imp	max-imp
CC (32)	8.832 (9.95)	0.325 (0.601)	0.848	1.02	34.69	0	2.227
OO (14)	6.778 (6.388)	0.314 (0.595)	0.513	1.02	21.43	0	2.24
OC-CO (19)	8.216 (8.089)	0.16 (0.16)	0.725	1.02	27.55	0	0.565
OT (3)	29.93 (48.318)	4.95 (8.314)	1	1.02	85.71	0.005	14.549

Table 8

Metrics of annotations ($n = 68$) in each domain and range category

R	ON	PON	OC	POC	MAI	IMP
R19	1	0.01	40	0.41	0.53	0.60
R18	1	0.01	49	0.50	0.36	0.56
R10	1	0.01	39	0.40	0.46	0.55
R1	1	0.01	34	0.35	0.47	0.52
R75	84	0.86	17	0.17	0.12	0.51
R27	1	0.01	38	0.39	0.35	0.48
R6	1	0.01	32	0.33	0.30	0.40
R7	2	0.02	36	0.37	0.24	0.40
R9	1	0.01	32	0.33	0.25	0.37
R35	1	0.01	34	0.35	0.19	0.34
R33	1	0.01	28	0.29	0.24	0.33
R13	2	0.02	33	0.34	0.18	0.33
R74	26	0.27	18	0.18	0.17	0.32
R15	2	0.02	31	0.32	0.14	0.29
R58	30	0.31	18	0.18	0.11	0.29

Table 9

The Top 15 requirement sets ordered by the their importance (IMP). ON: Number of ontologies for which requirement set is necessary. PON: ON as proportion. OC: Number of ontologies which are completely covered by requirement set. POC: OC as proportion. MAI: Mean importance of annotations in requirement set. IMP: Overall importance of requirement set. Shaded in grey or those requirements which are on the Pareto frontier w.r.t to PON, POC and IMA.

necessity score and the coverage: if the coverage is high, requirements of many ontologies are fully met (i.e. typically larger requirement sets), but conversely, not many ontologies strictly need (necessity score) the whole set of requirements to be fulfilled.

The average number of annotations per requirement is 7.733 ($\sigma = 6.831$), and ranges from 1 to 39. The top 15 requirements (w.r.t importance) have an average of 15.8 ($\sigma = 10.339$) annotations per requirement, over double the average score for all requirements. The necessity scores range from 1 to 84, and on average, each requirement set is necessarily needed for 7 ontologies. The coverage scores range from 1 to 49, and on average, 21 ontologies are completely covered per requirement.

When considering the diversity of annotations within each requirement set, on average, 44% of annotations are from the CC category (relations between continuants, e.g., *contains*), 15% from the OO category (relations between occurrents, e.g., *precedes*), 23% from the OC-CO category (relations between occurrents and continuants, e.g., *existenceStartsDuring*) and 16% are from the OT category (e.g., *partOf*). The annotation that occurs most often is A68, which occurs in 61 out of 75 (81%) requirements and annotates the *partOf* and *hasPart* relations.

Out of those 14 requirements on the Pareto frontier, A68 occurs in 13 requirements. The diversity of the 14 requirements is as follows: on average, 42% of the requirement sets' annotations are from the CC category, 14% from the OO category, 6% from the OC-CO category and 33% are from the OT category.

Considering only the top 3 requirement sets, the diversity of annotations along with their attributes remains high. 49 annotations are used within the top 3 requirement sets made up of 64 attributes. 22 of the annotations belong to the CC category, 10 to OO, 15 to OC-CO and 2 to OT. 28 of their attributes belong to the CC category, 12 to OO, 5 to OT and 19 to OC-CO. The diversity within each domain category is also high. For example, regarding the CC category which contains 28 attributes, 9 of these attributes come from the ***States*** set, 5 from the ***Time*** set, 9 from the ***Domain & Range*** set, 2 from the ***Identity*** set, 2 from the ***Rigid*** set and 1 from the ***AHFAT*** set. Similar characteristics can be seen in the remaining categories. Only 9 requirements (R2, R5, R42, R63, R66, R68, R69, R72 and R75) have annotations from only one domain and range category. Only one of these requirement sets (R75) falls outside the bottom 15 requirements sets ordered by importance. 20 requirement sets have annotations from 2 categories, another 23 have annotations from 3 categories and the remaining 23 requirement sets contain annotations from all 4 categories.

This demonstrates the level of coverage needed by a suitable temporal language extension to OWL. Based on the requirement sets, it would not be enough for a language extension to only focus on one type of temporal phenomena (for example, the modelling of continuants) as a wide variety of temporal phenomena exist in single requirements alone.

5. Discussion

To the best of our knowledge, this is the first study to systematically assess and report on a set of requirements for ontologies in a particular domain. By using a temporally annotated data set that is used widely across the ontology corpus, we were able to determine which individual temporal features in the dataset are most important, as well as their co-occurrence with other temporal features, both in terms of their usage in each ontology, and their coverage.

When considering the individual temporal features, due to the extent of diversity between the features, they were analysed in groups, categorised by their occurrence with the different domain and range features. We found that certain attributes were more prominent in the corpus than others. For example, when considering temporal features belonging to the CC category (those features used in relations who's domain and range type were both continuants), same-time relations were more common than both past-time and future-time relations. Due to the nature of the encoding scheme, we were also able to compare relation categories against each other. OT relations were overall the most prominent amongst the corpus (in terms of coverage and impact), followed by CC relations. OO and OC-CO relations had roughly the same usage.

The analysis of the defined requirements showed that there is high diversity amongst ontologies w.r.t the different categories of temporal phenomena. On average, we found that requirements are made up of just under half of CC features, followed by a quarter of OC-CO features, and the rest are made up OT and OO features. However, when focussing on the most important requirements, this shifts slightly in favour of the OT features. This is an important result since it shows that in order to meet the requirements, a language would have to be able to model a large set of temporal features. This may be difficult due to how different the features are in nature. For example, being able to model both continuants and occurrents may be difficult, due to how temporally different these entities are.

Amongst all stages of analysis, the relations *part of* and *has part*, along with their annotation, features and presence in requirements, were considered most important. These relations were the most used relations, both in terms of coverage and impact. Their features and annotation had the highest scores for coverage and impact, and their annotation was used in 81% of all requirements, and 92% of the most important requirements, and 100% of the requirements on the Pareto frontier. Arguably, the most interesting feature of these relations were the rigid attributes. It is well known that having the ability to model rigidity in temporal logics is a computationally hard problem [11, 17], which often leads to undecidability. If this is considered to be one of the most important features, many potential temporal language candidates may be deemed unsuitable.

Although not studied in detail in this paper, the analyses of the data and the definition of the requirements are intended to aid in the identification of a suitable temporal extension of OWL (or its underlying logic) to better aid in the modelling of the temporal features found. We showed that the level of coverage needed for even single requirements was very high. Language designers can use the requirement sets to determine how effective their languages are and to determine how best to extend their language if it is not suitable. They could also be used to drive the development of new language extensions based solely on the requirements found in this study. Languages could also be compared based on how many temporal requirements are met.

5.1. Limitations

Although we identified a large amount of temporal features present in the corpus of ontologies, they do not represent an exhaustive set of features. All features used were only derived from the relations used in RO. Ontologies may exhibit other types of temporal phenomena outside of the relation space which was not covered by this survey. Therefore, we can only claim to have defined a subset of the temporal requirements of the ontologies. At the present time, it is not clear how additional data could be extracted in a systematic or automated way, not only due to the size of ontologies and the additional time needed for manual inspection, but also due to there not being another known shared resource such as the Relation Ontology, or the Basic Formal Ontology, allowing data to be easily analysed.

When running our survey, we relied heavily on the notion of *smart matching*: a way to match relations across terminologies that *look* similar, but use different IRIs. Although our matching technique was sensible, it is possible that some of the matches may have been incorrect, or other matches may have been missed. Manual inspection of a sample of the matched relations suggested otherwise, however, some matches could still be missed.

5.2. Outlook

Before beginning to evaluate temporal language extensions, our next steps include further verification of our requirement results. We hope to achieve this by contacting ontology authors and confirming (1) whether our interpretation of their ontology's requirements was correct (2) whether our smart matching results were valid, and (3) whether our temporal interpretations of relations coincide with their own interpretations. This would reinforce the validity of our results and possibly make them more fine-grained: determining how relations are intended to be interpreted on an individual ontology level would allow us to eliminate the *Necessary* attributes (e.g. Rigid:Yes-Necessary:No), which would eliminate uncertainty in the requirements.

The system we created for defining the importance of certain features used throughout ontologies could be used in other application domains to determine importance of entities, not necessarily temporal. We intend to further generalise this procedure and apply it to other application domains to test its efficacy as an entity importance measuring system for ontologies.

6. Conclusion

Our study produced an empirically validated set of requirements that describe the temporal content of ontologies in the bio-health domain. The results showed that the temporal requirements are diverse and cover a wide range of different phenomena. These results aim to provide a mechanism to show which temporal language extensions are most suitable for the temporal modelling of bio-health ontologies and can also drive the creation of new language extensions, specifically tailored to the requirements and the temporal nature of bio-health ontologies.

References

[1] B.C. Grau, I. Horrocks, B. Motik, B. Parsia, P.F. Patel-Schneider and U. Sattler, OWL 2: The Next Step for OWL, *Journal of Web Semantics* **6**(4) (2008), 309–322.

[2] F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi and P.F. Patel-Schneider (eds), The Description Logic Handbook: Theory, Implementation, and Applications, in: *Description Logic Handbook*, Cambridge University Press. ISBN 0-521-78176-0.

[3] F. Baader and U. Sattler, An Overview of Tableau Algorithms for Description Logics, *Studia Logica* **69**(1) (2001), 5–40.

[4] M. Krötzsch, F. Simancik and I. Horrocks, A Description Logic Primer, *CoRR* **abs/1201.4089** (2012).

[5] I. Horrocks, O. Kutz and U. Sattler, The Even More Irresistible SROIQ, in: *KR*, 2006, pp. 57–67.

[6] S. Brandt, Reasoning in ELH w.r.t. General Concept Inclusion Axioms, Technical Report, 2004.

[7] F. Baader, S. Brand and C. Lutz, Pushing the EL Envelope, in: *Proc. of IJCAI 2005*, Morgan-Kaufmann Publishers, 2005, pp. 364–369.

[8] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L.J. Goldberg, K. Eilbeck, A. Ireland, C.J. Mungall, OBI Consortium, N. Leontis, P. Rocca-Serra, A. Ruttenberg, S.-A.A. Sansone, R.H. Scheuermann, N. Shah, P.L. Whetzel and S. Lewis, The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration., *Nature biotechnology* **25**(11) (2007), 1251–1255, ISSN 1087-0156. doi:10.1038/nbt1346. http://dx.doi.org/10.1038/nbt1346.

[9] M. Costa, S. Reeve, G. Grumbling and D. Osumi-Sutherland, The Drosophila Anatomy Ontology, *Journal of Biomedical Semantics* **4**(1) (2013), 32, ISSN 2041-1480. doi:10.1186/2041-1480-4-32. http://www.jbiomedsem.com/content/4/1/32.

[10] N. Sioutos, S. de Coronado, M.W. Haber, F.W. Hartel, W. Shaiu and L.W. Wright, NCI Thesaurus: A Semantic Model Integrating Cancer-related Clinical and Molecular Information, *Journal of Biomedical Informatics* **40**(1) (2007), 30–43. doi:10.1016/j.jbi.2006.02.013. http://dx.doi.org/10.1016/j.jbi.2006.02.013.

[11] C. Lutz, F. Wolter and M. Zakharyaschev, Temporal Description Logics: A Survey, in: *TIME*, S. Demri and C.S. Jensen, eds, IEEE Computer Society, 2008, pp. 3–14. ISBN 978-0-7695-3181-6.

[12] V. Gutiérrez-Basulto, J.C. Jung and T. Schneider, The Complexity of Temporal Description Logics with Rigid Roles and Restricted TBoxes: In Quest of Saving a Troublesome Marriage, in: *Proceedings of the 28th International Workshop on Description Logics, Athens,Greece, June 7-10, 2015.*, D. Calvanese and B. Konev, eds, CEUR Workshop Proceedings, Vol. 1350, CEUR-WS.org, 2015. http://ceur-ws.org/Vol-1350/paper-23.pdf.

[13] F. Baader and P. Hanschke, A Scheme for Integrating Concrete Domains into Concept Languages, in: *IJCAI*, J. Mylopoulos and R. Reiter, eds, Morgan Kaufmann, 1991, pp. 452–457. ISBN 1-55860-160-0. http://ijcai.org/proceedings/1991-1.

[14] V. Milea, F. Frasincar and U. Kaymak, tOWL: A Temporal Web Ontology Language.

[15] C.A. Welty and R. Fikes, A Reusable Ontology for Fluents in OWL, in: *Formal Ontology in Information Systems, Proceedings of the Fourth International Conference, FOIS 2006, Baltimore, Maryland, USA, November 9-11, 2006*, 2006, pp. 226–236. http://www.booksonline.iospress.nl/Content/View.aspx?piid=2209.

[16] C. Little and S. Cox, Time Ontology in OWL (2017), https://www.w3.org/TR/2017/REC-owl-time-20171019/.

[17] C. Lutz, Description Logics with Concrete Domains-A Survey, in: *Advances in Modal Logic*, P. Balbiani, N. Suzuki, F. Wolter and M. Zakharyaschev, eds, King's College Publications, 2002, pp. 265–296. ISBN 0-9543006-2-9.

[18] P. Grenon, B. Smith and L.J. Goldberg, Biodynamic ontology: applying BFO in the biomedical domain., *Studies in health technology and informatics* **102** (2004), 20–38.

[19] B. Smith, W. Ceusters, B. Klagges, J. Köhler, A. Kuma, J. Lomax, C. Mungall, F. Neuhaus, A. Rector and C. Rosse, Relations in Biomedical Ontologies, *Genome Biology* **6**(5) (2005), 46.

[20] J.F. Allen, Maintaining Knowledge about Temporal Intervals, *Commun. ACM* **26**(11) (1983), 832–843, ISSN 0001-0782. doi:10.1145/182.358434. http://doi.acm.org/10.1145/182.358434.

[21] S. Demri and C.S. Jensen (eds), 15th International Symposium on Temporal Representation and Reasoning, TIME 2008, Université du Québec à Montréal, Canada, 16-18 June 2008, in: *TIME*, IEEE Computer Society, 2008. ISBN 978-0-7695-3181-6.

[22] J. Mylopoulos and R. Reiter (eds), Proceedings of the 12th International Joint Conference on Artificial Intelligence. Sydney, Australia, August 24-30, 1991, Morgan Kaufmann, 1991. ISBN 1-55860-160-0. http://ijcai.org/proceedings/1991-1.

[23] P. Balbiani, N. Suzuki, F. Wolter and M. Zakharyaschev (eds), Advances in Modal Logic 4, papers from the fourth conference on "Advances in Modal logic," held in Toulouse (France) in October 2002, King's College Publications, 2003. ISBN 0-9543006-2-9.

[24] D. Calvanese and B. Konev (eds), Proceedings of the 28th International Workshop on Description Logics, Athens,Greece, June 7-10, 2015, in *CEUR Workshop Proceedings*, Vol. 1350, CEUR-WS.org, 2015. http://ceur-ws.org/Vol-1350.

Appendix A. Relations

A.1. Relations: Coverage

Relation	**#O**	**COV**	**CAT**
inheres in	29	29.59	CC
has role	26	26.53	CC
located in	21	21.43	CC
has quality	20	20.41	CC
bearer of	19	19.39	CC
develops from	19	19.39	CC
derives from	16	16.33	CC
adjacent to	15	15.31	CC
concretizes	15	15.31	CC
has function	10	10.20	CC
has member	10	10.20	CC
towards	10	10.20	CC
overlaps	9	9.18	CC
continuous with	8	8.16	CC
composed primarily of	7	7.14	CC
has component	7	7.14	CC
location of	7	7.14	CC
member of	7	7.14	CC
surrounded by	7	7.14	CC
function of	6	6.12	CC
is concretized as	6	6.12	CC
produces	6	6.12	CC
role of	6	6.12	CC
surrounds	6	6.12	CC
attached to	5	5.10	CC
inheres in part of	5	5.10	CC
connected to	4	4.08	CC
connects	4	4.08	CC
has developmental contribution from	4	4.08	CC
innervates	4	4.08	CC
produced by	4	4.08	CC
bounding layer of	3	3.06	CC
contains	3	3.06	CC
develops into	3	3.06	CC
directly develops from	3	3.06	CC
has potential to develop into	3	3.06	CC
innervated_by	3	3.06	CC
conduit for	2	2.04	CC
contributes to morphology of	2	2.04	CC
developmentally induced by	2	2.04	CC
developmentally replaces	2	2.04	CC
develops in	2	2.04	CC
has 2D boundary	2	2.04	CC
has habitat	2	2.04	CC
has modifier	2	2.04	CC
has plasma membrane part	2	2.04	CC
has potential to developmentally contribute to	2	2.04	CC
has skeleton	2	2.04	CC
has soma location	2	2.04	CC
has synaptic terminal in	2	2.04	CC
immediate transformation of	2	2.04	CC
interacts with	2	2.04	CC
luminal space of	2	2.04	CC
quality of	2	2.04	CC
skeleton of	2	2.04	CC
supplies	2	2.04	CC
synapsed by	2	2.04	CC
synapsed to	2	2.04	CC
transformation of	2	2.04	CC
tributary of	2	2.04	CC
attached to part of	1	1.02	CC
branching part of	1	1.02	CC

Relation	#O	COV	CAT
child nucleus of	1	1.02	CC
child nucleus of in hermaphrodite	1	1.02	CC
child nucleus of in male	1	1.02	CC
confers advantage in	1	1.02	CC
contained in	1	1.02	CC
determined by	1	1.02	CC
determined by part of	1	1.02	CC
develops from part of	1	1.02	CC
distributary of	1	1.02	CC
drains	1	1.02	CC
electrically_synapsed_to	1	1.02	CC
expresses	1	1.02	CC
fasciculates with	1	1.02	CC
gene product of	1	1.02	CC
has disposition	1	1.02	CC
has fused element	1	1.02	CC
has host	1	1.02	CC
has muscle antagonist	1	1.02	CC
has muscle insertion	1	1.02	CC
has muscle origin	1	1.02	CC
has postsynaptic terminal in	1	1.02	CC
has presynaptic terminal in	1	1.02	CC
has synaptic terminal of	1	1.02	CC
has vector	1	1.02	CC
in homology relationship with	1	1.02	CC
lumen of	1	1.02	CC
molecularly interacts with	1	1.02	CC
partially overlaps	1	1.02	CC
serially homologous to	1	1.02	CC
spatially disjoint from	1	1.02	CC
synapsed_via_type_Ib_bouton_to	1	1.02	CC
synapsed_via_type_II_bouton_to	1	1.02	CC
synapsed_via_type_III_bouton_to	1	1.02	CC
synapsed_via_type_Is_bouton_to	1	1.02	CC
transcribed from	1	1.02	CC
transcribed to	1	1.02	CC
has participant	27	27.55	OC-CO
realizes	24	24.49	OC-CO
realized in	17	17.35	OC-CO
participates in	15	15.31	OC-CO
occurs in	14	14.29	OC-CO
capable of	10	10.20	OC-CO

Relation	#O	COV	CAT
has output	8	8.16	OC-CO
output of	6	6.12	OC-CO
has input	5	5.10	OC-CO
existence starts during	4	4.08	OC-CO
existence starts during or after	4	4.08	OC-CO
capable of part of	3	3.06	OC-CO
existence ends during	3	3.06	OC-CO
existence ends during or before	3	3.06	OC-CO
existence starts and ends during	3	3.06	OC-CO
actively participates in	2	2.04	OC-CO
existence ends with	2	2.04	OC-CO
existence starts with	2	2.04	OC-CO
formed as result of	2	2.04	OC-CO
has active participant	2	2.04	OC-CO
results in formation of	2	2.04	OC-CO
contains process	1	1.02	OC-CO
functionally related to	1	1.02	OC-CO
has intermediate	1	1.02	OC-CO
preceded by	18	18.37	OO
immediately preceded by	15	15.31	OO
precedes	15	15.31	OO
regulates	10	10.20	OO
negatively regulates	6	6.12	OO
starts	6	6.12	OO
ends during	4	4.08	OO
positively regulates	4	4.08	OO
ends	3	3.06	OO
happens during	3	3.06	OO
obsolete preceded by	3	3.06	OO
ends with	2	2.04	OO
immediately precedes	2	2.04	OO
starts during	2	2.04	OO
starts with	2	2.04	OO
causally downstream of	1	1.02	OO
causally upstream of or within	1	1.02	OO
simultaneous with	1	1.02	OO
part of	78	79.59	OT
has part	48	48.98	OT
in taxon	2	2.04	OT
only in taxon	2	2.04	OT
depends on	1	1.02	OT

Relation	**#O**	**COV**	**CAT**

Table 12: Temporal relations, grouped by temporal category and ordered by coverage (COV).

A.2. Relations: Impact

RO Relation	**IMP**	**CAT**
inheres in	2.066	CC
has quality	1.521	CC
bearer of	1.302	CC
develops from	0.994	CC
has modifier	0.651	CC
derives from	0.571	CC
has role	0.530	CC
overlaps	0.341	CC
has component	0.206	CC
attached to	0.194	CC
concretizes	0.158	CC
has function	0.147	CC
towards	0.132	CC
has member	0.131	CC
has plasma membrane part	0.129	CC
child nucleus of	0.126	CC
inheres in part of	0.123	CC
located in	0.115	CC
composed primarily of	0.092	CC
innervated_by	0.076	CC
directly develops from	0.048	CC
adjacent to	0.048	CC
continuous with	0.045	CC
has postsynaptic terminal in	0.038	CC
gene product of	0.037	CC
has presynaptic terminal in	0.035	CC
child nucleus of in male	0.035	CC
has disposition	0.028	CC
has synaptic terminal in	0.026	CC
child nucleus of in hermaphrodite	0.026	CC
interacts with	0.023	CC
is concretized as	0.022	CC
surrounds	0.022	CC
role of	0.021	CC
has soma location	0.020	CC
connected to	0.020	CC
fasciculates with	0.018	CC
has potential to develop into	0.017	CC
expresses	0.017	CC
contributes to morphology of	0.017	CC
has developmental contribution from	0.016	CC
produces	0.016	CC
connects	0.015	CC
synapsed by	0.012	CC
quality of	0.010	CC
member of	0.010	CC
location of	0.010	CC
synapsed to	0.010	CC
has host	0.008	CC
surrounded by	0.008	CC
innervates	0.007	CC
has skeleton	0.007	CC
skeleton of	0.007	CC
immediate transformation of	0.007	CC
spatially disjoint from	0.006	CC
in homology relationship with	0.005	CC
develops into	0.005	CC
function of	0.005	CC
has muscle insertion	0.005	CC
has vector	0.005	CC
transcribed to	0.005	CC
branching part of	0.004	CC
has muscle origin	0.004	CC
has potential to developmentally contribute to	0.004	CC
produced by	0.004	CC
luminal space of	0.003	CC
synapsed_via_type_Ib_bouton_to	0.003	CC
bounding layer of	0.003	CC
develops in	0.003	CC
contains	0.003	CC
supplies	0.003	CC
tributary of	0.002	CC
determined by	0.002	CC
molecularly interacts with	0.002	CC
conduit for	0.002	CC
drains	0.001	CC
has synaptic terminal of	0.001	CC
developmentally induced by	0.001	CC
transformation of	0.001	CC
has fused element	0.001	CC
has muscle antagonist	0.001	CC
determined by part of	0.001	CC

RO Relation	IMP	CAT
developmentally replaces	0.001	CC
electrically_synapsed_to	0.001	CC
transcribed from	0.001	CC
has habitat	0.001	CC
distributary of	0.000	CC
has 2D boundary	0.000	CC
synapsed_via_type_II_bouton_to	0.000	CC
synapsed_via_type_Is_bouton_to	0.000	CC
attached to part of	0.000	CC
confers advantage in	0.000	CC
develops from part of	0.000	CC
partially overlaps	0.000	CC
serially homologous to	0.000	CC
synapsed_via_type_III_bouton_to	0.000	CC
contained in	0.000	CC
lumen of	0.000	CC
has participant	0.560	OC-CO
realized in	0.385	OC-CO
participates in	0.266	OC-CO
existence ends during	0.263	OC-CO
existence starts during or after	0.261	OC-CO
existence starts during	0.260	OC-CO
existence ends during or before	0.259	OC-CO
realizes	0.233	OC-CO
occurs in	0.228	OC-CO
capable of	0.145	OC-CO
has output	0.063	OC-CO
has input	0.029	OC-CO
output of	0.025	OC-CO
existence starts and ends during	0.022	OC-CO
formed as result of	0.017	OC-CO
has active participant	0.005	OC-CO
capable of part of	0.005	OC-CO
actively participates in	0.004	OC-CO
results in formation of	0.004	OC-CO
has intermediate	0.002	OC-CO
existence ends with	0.001	OC-CO
existence starts with	0.001	OC-CO
contains process	0.001	OC-CO
functionally related to	0.000	OC-CO
immediately preceded by	2.240	OO
preceded by	0.560	OO
ends during	0.496	OO
starts during	0.496	OO
regulates	0.157	OO
happens during	0.116	OO
negatively regulates	0.083	OO
precedes	0.072	OO
positively regulates	0.071	OO
obsolete preceded by	0.060	OO
immediately precedes	0.023	OO
causally downstream of	0.012	OO
starts	0.007	OO
ends	0.000	OO
ends with	0.000	OO
simultaneous with	0.000	OO
starts with	0.000	OO
causally upstream of or within	0.000	OO
part of	11.519	OT
has part	3.029	OT
in taxon	0.276	OT
only in taxon	0.020	OT
depends on	0.005	OT

Table 13: Temporal relations, grouped by temporal category and ordered by impact (IMP).

Appendix B. Temporal Attributes

B.1. Temporal Attributes: Coverage

Temporal Attribute	#O	COV	CAT
CC-Dom:C-Ran:C	68	69.39	CC
CC-Dom:IC-Ran:C	62	63.27	CC
CC-Time:Same	60	61.22	CC
CC-Rig:Yes-Nec:No	59	60.20	CC
CC-TI:AHFAT	53	54.08	CC
CC-Dom:C-Ran:IC	46	46.94	CC
CC-Dom:IC-Ran:IC	46	46.94	CC
CC-Dom:IC-Ran:SDC	38	38.78	CC
CC-Time:Diff	37	37.76	CC
CC-Time:Past	36	36.73	CC
CC-Dom:SDC-Ran:C	34	34.69	CC
CC-Dom:SDC-Ran:IC	30	30.61	CC
CC-Dom:Birth	22	22.45	CC

Temporal Attribute	#O	COV	CAT
CC-Dom:Changed	22	22.45	CC
CC-Ran:Changed	22	22.45	CC
CC-Dom:Birth-Nec:No	19	19.39	CC
CC-Dom:Changed-Nec:No	19	19.39	CC
CC-Identity:Same-Nec:No	19	19.39	CC
CC-Ran:Death	17	17.35	CC
CC-Dom:SDC-Ran:GDC	15	15.31	CC
CC-Time:Future	9	9.18	CC
CC-Ran:Birth	7	7.14	CC
CC-Dom:GDC-Ran:SDC	6	6.12	CC
CC-Rig:Yes	6	6.12	CC
CC-Ran:Changed-Nec:No	5	5.10	CC
CC-Ran:Birth-Nec:No	3	3.06	CC
CC-Dom:SDC-Ran:SDC	2	2.04	CC
CC-Identity:Same	2	2.04	CC
CC-Ran:Death-Nec:No	2	2.04	CC
CC-Time:PastImmediate	2	2.04	CC
CC-Dom:GDC-Ran:GDC	1	1.02	CC
OC-CO-Time:Same	46	46.94	OC-CO
OC-CO-Dom:O-Ran:C	42	42.86	OC-CO
OC-CO-Dom:P-Ran:C	37	37.76	OC-CO
OC-CO-Rig:Yes-Nec:No	36	36.73	OC-CO
OC-CO-Dom:C-Ran:O	34	34.69	OC-CO
OC-CO-Dom:C-Ran:P	33	33.67	OC-CO
OC-CO-Dom:P-Ran:SDC	24	24.49	OC-CO
OC-CO-Rig:Yes	18	18.37	OC-CO
OC-CO-Dom:SDC-Ran:P	17	17.35	OC-CO
OC-CO-Dom:O-Ran:IC	14	14.29	OC-CO
OC-CO-Time:Diff	13	13.27	OC-CO
OC-CO-Dom:IC-Ran:O	11	11.22	OC-CO
OC-CO-Dom:IC-Ran:P	11	11.22	OC-CO
OC-CO-Time:Future	10	10.20	OC-CO
OC-CO-Ran:Birth	9	9.18	OC-CO
OC-CO-Ran:Changed	9	9.18	OC-CO
OC-CO-Dom:Birth	7	7.14	OC-CO
OC-CO-Dom:Changed	7	7.14	OC-CO
OC-CO-Dom:Death	4	4.08	OC-CO
OC-CO-Time:Same/Past	4	4.08	OC-CO
OC-CO-Time:Same/Future	3	3.06	OC-CO

Temporal Attribute	#O	COV	CAT
OO-Dom:O-Ran:O	41	41.84	OO
OO-Time:All	41	41.84	OO
OO-Time:BeforeInverse	21	21.43	OO
OO-Time:Before	15	15.31	OO
OO-Time:MeetsInverse	15	15.31	OO
OO-Time:Before/During	13	13.27	OO
OO-Dom:P-Ran:P	10	10.20	OO
OO-Time:Starts	6	6.12	OO
OO-Time:During/Overlaps	4	4.08	OO
OO-Time:During	3	3.06	OO
OO-Time:Finishes	3	3.06	OO
OO-Time:During/OverlapsInverse	2	2.04	OO
OO-Time:FinishesInverse	2	2.04	OO
OO-Time:Meets	2	2.04	OO
OO-Time:StartsInverse	2	2.04	OO
OO-Time:IsEqualTo	1	1.02	OO
OT-Dom:X-Ran:X	84	85.71	OT
OT-Rig:Yes-Nec:No	84	85.71	OT
OT-TI:AHFAT	84	85.71	OT
OT-Time:Same	84	85.71	OT
OT-Dom:X-Ran:IC	3	3.06	OT

Table 14: Temporal attributes, grouped by temporal category and ordered by coverage (COV).

B.2. Temporal Attributes: Impact

Temporal Attribute	IMP	CAT
CC-Dom:C-Ran:C	10.398	CC
CC-Time:Same	8.488	CC
CC-Rig:Yes-Nec:No	8.221	CC
CC-Dom:IC-Ran:C	6.721	CC
CC-TI:AHFAT	4.827	CC
CC-Dom:IC-Ran:SDC	3.529	CC
CC-Dom:SDC-Ran:C	3.166	CC
CC-Dom:C-Ran:IC	3.114	CC
CC-Dom:IC-Ran:IC	3.112	CC
CC-Dom:SDC-Ran:IC	2.225	CC
CC-Time:Diff	1.910	CC
CC-Time:Past	1.867	CC
CC-Identity:Same-Nec:No	1.065	CC
CC-Dom:Birth-Nec:No	1.060	CC
CC-Dom:Changed-Nec:No	1.060	CC

Temporal Attribute	IMP	CAT
CC-Dom:Changed	0.808	CC
CC-Dom:Birth	0.798	CC
CC-Ran:Changed	0.778	CC
CC-Ran:Death	0.757	CC
CC-Dom:SDC-Ran:SDC	0.651	CC
CC-Dom:SDC-Ran:GDC	0.158	CC
CC-Rig:Yes	0.091	CC
CC-Time:Future	0.043	CC
CC-Dom:GDC-Ran:SDC	0.022	CC
CC-Ran:Birth	0.021	CC
CC-Identity:Same	0.008	CC
CC-Time:PastImmediate	0.007	CC
CC-Ran:Changed-Nec:No	0.006	CC
CC-Dom:GDC-Ran:GDC	0.005	CC
CC-Ran:Birth-Nec:No	0.005	CC
CC-Ran:Death-Nec:No	0.001	CC
OC-CO-Time:Same	2.376	OC-CO
OC-CO-Dom:C-Ran:O	1.916	OC-CO
OC-CO-Rig:Yes-Nec:No	1.454	OC-CO
OC-CO-Dom:O-Ran:C	1.124	OC-CO
OC-CO-Dom:Changed	1.084	OC-CO
OC-CO-Dom:P-Ran:C	0.896	OC-CO
OC-CO-Dom:C-Ran:P	0.849	OC-CO
OC-CO-Time:Diff	0.664	OC-CO
OC-CO-Dom:Birth	0.561	OC-CO
OC-CO-Dom:Death	0.545	OC-CO
OC-CO-Rig:Yes	0.517	OC-CO
OC-CO-Dom:SDC-Ran:P	0.385	OC-CO
OC-CO-Time:Same/Past	0.261	OC-CO
OC-CO-Time:Same/Future	0.259	OC-CO
OC-CO-Dom:P-Ran:SDC	0.233	OC-CO
OC-CO-Dom:O-Ran:IC	0.228	OC-CO
OC-CO-Dom:IC-Ran:O	0.151	OC-CO
OC-CO-Dom:IC-Ran:P	0.151	OC-CO
OC-CO-Time:Future	0.145	OC-CO
OC-CO-Ran:Changed	0.096	OC-CO
OC-CO-Ran:Birth	0.066	OC-CO
OO-Dom:O-Ran:O	4.393	OO
OO-Time:All	4.393	OO
OO-Time:MeetsInverse	2.240	OO
OO-Time:BeforeInverse	0.632	OO
OO-Time:During/Overlaps	0.496	OO
OO-Time:During/OverlapsInverse	0.496	OO
OO-Time:Before/During	0.311	OO
OO-Dom:P-Ran:P	0.157	OO
OO-Time:During	0.116	OO
OO-Time:Before	0.072	OO
OO-Time:Meets	0.023	OO
OO-Time:Starts	0.007	OO
OO-Time:Finishes	0.000	OO
OO-Time:FinishesInverse	0.000	OO
OO-Time:IsEqualTo	0.000	OO
OO-Time:StartsInverse	0.000	OO
OT-Time:Same	14.849	OT
OT-Dom:X-Ran:X	14.553	OT
OT-Rig:Yes-Nec:No	14.553	OT
OT-TI:AHFAT	14.549	OT
OT-Dom:X-Ran:IC	0.296	OT

Table 15: Temporal attributes, grouped by temporal category and ordered by impact (IMP).

Appendix C. Annotations

C.1. Temporal attributes of annotations

ID	Attributes	Inferred Attributes
A1	Dom:C-Ran:C Time:Same	
A2	Dom:C-Ran:C Time:Same Rig:Yes-Nec:No TI:AHFAT	
A3	Dom:C-Ran:IC Time:Same Rig:Yes-Nec:No TI:AHFAT	Dom:C-Ran:C
A4	Dom:C-Ran:O Time:Same/Future Dom:Death	Time:Diff Dom:Changed
A5	Dom:C-Ran:O Time:Same/Past Dom:Birth	Time:Diff Dom:Changed
A6	Dom:C-Ran:O Time:Same Dom:Birth	Dom:Changed

ID	Attributes	Inferred Attributes
A7	Dom:C-Ran:O Time:Same Rig:Yes Dom:Birth Dom:Death	Dom:Changed
A8	Dom:C-Ran:O Time:Same Rig:Yes Dom:Death	Dom:Changed
A9	Dom:C-Ran:P Time:Same	Dom:C-Ran:O
A10	Dom:C-Ran:P Time:Same Dom:Birth	Dom:C-Ran:O Dom:Changed
A11	Dom:C-Ran:P Time:Same Rig:Yes-Nec:No	Dom:C-Ran:O
A12	Dom:GDC-Ran:GDC Time:Future Ran:Birth	Time:Diff Ran:Changed Dom:C-Ran:C
A13	Dom:GDC-Ran:GDC Time:Past Dom:Birth	Time:Diff Dom:C-Ran:C Dom:Changed
A14	Dom:GDC-Ran:SDC Time:Same Rig:Yes-Nec:No	Dom:C-Ran:C
A15	Dom:IC-Ran:C Time:Past Dom:Birth	Time:Diff Dom:Changed Dom:C-Ran:C
A16	Dom:IC-Ran:C Time:Same	Dom:C-Ran:C
A17	Dom:IC-Ran:C Time:Same Rig:Yes-Nec:No TI:AHFAT	Dom:C-Ran:C
A18	Dom:IC-Ran:IC Time:Future Identity:Same-Nec:No Dom:Birth-Nec:No	Time:Diff Dom:Changed-Nec:No Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A19	Dom:IC-Ran:IC Time:Future Identity:Same-Nec:No Ran:Birth-Nec:No	Time:Diff Ran:Changed-Nec:No Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A20	Dom:IC-Ran:IC Time:Future Ran:Birth	Time:Diff Dom:C-Ran:IC Ran:Changed Dom:C-Ran:C Dom:IC-Ran:C
A21	Dom:IC-Ran:IC Time:Past	Time:Diff Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A22	Dom:IC-Ran:IC Time:Past Dom:Birth	Time:Diff Dom:C-Ran:IC Dom:C-Ran:C Dom:Changed Dom:IC-Ran:C
A23	Dom:IC-Ran:IC Time:Past Dom:Birth Ran:Death	Time:Diff Dom:C-Ran:IC Ran:Changed Dom:C-Ran:C Dom:Changed Dom:IC-Ran:C
A24	Dom:IC-Ran:IC Time:Past Dom:Changed	Time:Diff Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A25	Dom:IC-Ran:IC Time:Past Dom:Changed Ran:Changed Dom:Birth-Nec:No Ran:Death-Nec:No	Time:Diff Ran:Changed-Nec:No Dom:Changed-Nec:No Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A26	Dom:IC-Ran:IC Time:Past Identity:Same-Nec:No Dom:Birth-Nec:No	Time:Diff Dom:Changed-Nec:No Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A27	Dom:IC-Ran:IC Time:Past Identity:Same Dom:Changed	Time:Diff Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A28	Dom:IC-Ran:IC Time:PastImmediate Identity:Same Dom:Changed	Time:Diff Time:Past Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C

ID	Attributes	Inferred Attributes
A29	Dom:IC-Ran:IC Time:Same	Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A30	Dom:IC-Ran:IC Time:Same Rig:Yes	Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A31	Dom:IC-Ran:IC Time:Same Rig:Yes-Nec:No	Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A32	Dom:IC-Ran:IC Time:Same Rig:Yes-Nec:No TI:AHFAT	Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A33	Dom:IC-Ran:IC Time:Same Rig:Yes TI:AHFAT	Dom:C-Ran:IC Dom:C-Ran:C Dom:IC-Ran:C
A34	Dom:IC-Ran:O Time:Same Rig:Yes-Nec:No	Dom:C-Ran:O
A35	Dom:IC-Ran:P Time:Future	Time:Diff Dom:C-Ran:P Dom:C-Ran:O Dom:IC-Ran:O
A36	Dom:IC-Ran:P Time:Same	Dom:C-Ran:P Dom:C-Ran:O Dom:IC-Ran:O
A37	Dom:IC-Ran:P Time:Same Rig:Yes	Dom:C-Ran:P Dom:C-Ran:O Dom:IC-Ran:O
A38	Dom:IC-Ran:SDC Time:Same Rig:Yes-Nec:No	Dom:C-Ran:C Dom:IC-Ran:C
A39	Dom:IC-Ran:SDC Time:Same Rig:Yes-Nec:No TI:AHFAT	Dom:C-Ran:C Dom:IC-Ran:C
A40	Dom:O-Ran:IC Time:Same Rig:Yes	Dom:O-Ran:C
A41	Dom:O-Ran:O Time:Before	Time:All
A42	Dom:O-Ran:O Time:Before/During	Time:All
A43	Dom:O-Ran:O Time:BeforeInverse	Time:All
A44	Dom:O-Ran:O Time:During	Time:All
A45	Dom:O-Ran:O Time:During/Overlaps	Time:All
A46	Dom:O-Ran:O Time:During/OverlapsInverse	Time:All
A47	Dom:O-Ran:O Time:Finishes	Time:All
A48	Dom:O-Ran:O Time:FinishesInverse	Time:All
A49	Dom:O-Ran:O Time:IsEqualTo	Time:All
A50	Dom:O-Ran:O Time:Meets	Time:All
A51	Dom:O-Ran:O Time:MeetsInverse	Time:All
A52	Dom:O-Ran:O Time:Starts	Time:All
A53	Dom:O-Ran:O Time:StartsInverse	Time:All
A54	Dom:P-Ran:C Time:Same Ran:Birth	Dom:O-Ran:C Ran:Changed
A55	Dom:P-Ran:C Time:Same Ran:Changed	Dom:O-Ran:C
A56	Dom:P-Ran:C Time:Same Rig:Yes	Dom:O-Ran:C
A57	Dom:P-Ran:C Time:Same Rig:Yes-Nec:No	Dom:O-Ran:C
A58	Dom:P-Ran:P Time:Before/During	Dom:O-Ran:O Time:All
A59	Dom:P-Ran:SDC Time:Same Rig:Yes-Nec:No	Dom:O-Ran:C Dom:P-Ran:C
A60	Dom:SDC-Ran:C Time:Same	Dom:C-Ran:C
A61	Dom:SDC-Ran:GDC Time:Same Rig:Yes-Nec:No	Dom:SDC-Ran:C Dom:C-Ran:C
A62	Dom:SDC-Ran:IC Time:Same Rig:Yes-Nec:No	Dom:SDC-Ran:C Dom:C-Ran:C
A63	Dom:SDC-Ran:IC Time:Same Rig:Yes-Nec:No TI:AHFAT	Dom:SDC-Ran:C Dom:C-Ran:C
A64	Dom:SDC-Ran:P Time:Same Rig:Yes-Nec:No	Dom:C-Ran:P Dom:C-Ran:O

ID	Attributes	Inferred Attributes
A65	Dom:SDC-Ran:SDC Time:Same Rig:Yes-Nec:No	Dom:SDC-Ran:C Dom:C-Ran:C
A66	Dom:X-Ran:IC Time:Same	
A67	Dom:X-Ran:X Time:Same Rig:Yes-Nec:No	
A68	Dom:X-Ran:X Time:Same Rig:Yes-Nec:No TI:AHFAT	

Table 16: List of all temporal annotations with their corresponding temporal attributes (explicit and inferred).

C.2. Annotations: Coverage

Annotation	#O	COV	CAT
A32	34	34.69	CC
A38	34	34.69	CC
A63	29	29.59	CC
A2	19	19.39	CC
A26	19	19.39	CC
A39	19	19.39	CC
A23	17	17.35	CC
A61	15	15.31	CC
A31	14	14.29	CC
A60	10	10.20	CC
A62	9	9.18	CC
A14	6	6.12	CC
A20	6	6.12	CC
A29	5	5.10	CC
A33	5	5.10	CC
A21	4	4.08	CC
A22	4	4.08	CC
A30	4	4.08	CC
A18	3	3.06	CC
A19	3	3.06	CC
A17	2	2.04	CC
A24	2	2.04	CC
A25	2	2.04	CC
A27	2	2.04	CC
A28	2	2.04	CC
A65	2	2.04	CC
A1	1	1.02	CC
A3	1	1.02	CC
A12	1	1.02	CC
A13	1	1.02	CC
A15	1	1.02	CC
A16	1	1.02	CC
A57	27	27.55	OC-CO
A59	24	24.49	OC-CO
A64	17	17.35	OC-CO
A11	16	16.33	OC-CO
A40	14	14.29	OC-CO
A35	10	10.20	OC-CO
A54	9	9.18	OC-CO
A9	6	6.12	OC-CO
A55	5	5.10	OC-CO
A5	4	4.08	OC-CO
A6	4	4.08	OC-CO
A4	3	3.06	OC-CO
A7	3	3.06	OC-CO
A8	3	3.06	OC-CO
A36	3	3.06	OC-CO
A10	2	2.04	OC-CO
A34	1	1.02	OC-CO
A37	1	1.02	OC-CO
A56	1	1.02	OC-CO
A43	21	21.43	OO
A41	15	15.31	OO
A51	15	15.31	OO
A58	10	10.20	OO
A42	7	7.14	OO
A52	6	6.12	OO
A45	4	4.08	OO
A44	3	3.06	OO
A47	3	3.06	OO
A46	2	2.04	OO
A48	2	2.04	OO
A50	2	2.04	OO
A53	2	2.04	OO
A49	1	1.02	OO
A68	84	85.71	OT

Annotation	#O	COV	CAT
A66	3	3.06	OT
A67	1	1.02	OT

Table 17: Temporal annotations, grouped by temporal category and ordered by coverage (COV).

C.3. Annotations: Impact

Annotation	IMP	CAT
A38	2.227	CC
A63	2.189	CC
A39	1.302	CC
A26	1.043	CC
A32	0.813	CC
A23	0.757	CC
A65	0.651	CC
A2	0.482	CC
A31	0.314	CC
A61	0.158	CC
A60	0.132	CC
A30	0.078	CC
A15	0.037	CC
A62	0.036	CC
A29	0.026	CC
A17	0.026	CC
A14	0.022	CC
A18	0.017	CC
A16	0.017	CC
A21	0.016	CC
A20	0.016	CC
A33	0.013	CC
A28	0.007	CC
A19	0.005	CC
A12	0.005	CC
A22	0.004	CC
A3	0.001	CC
A24	0.001	CC
A27	0.001	CC
A13	0.001	CC
A25	0.001	CC
A1	0.000	CC
A57	0.565	OC-CO
A64	0.385	OC-CO
A11	0.271	OC-CO
A8	0.264	OC-CO
A6	0.262	OC-CO
A5	0.261	OC-CO
A4	0.259	OC-CO
A59	0.233	OC-CO
A40	0.228	OC-CO
A35	0.145	OC-CO
A54	0.066	OC-CO
A55	0.029	OC-CO
A9	0.025	OC-CO
A7	0.022	OC-CO
A10	0.017	OC-CO
A36	0.005	OC-CO
A56	0.002	OC-CO
A37	0.001	OC-CO
A34	0.000	OC-CO
A51	2.240	OO
A43	0.632	OO
A45	0.496	OO
A46	0.496	OO
A58	0.157	OO
A42	0.154	OO
A44	0.116	OO
A41	0.072	OO
A50	0.023	OO
A52	0.007	OO
A47	0.000	OO
A48	0.000	OO
A49	0.000	OO
A53	0.000	OO
A68	14.549	OT
A66	0.296	OT
A67	0.005	OT

Table 18: Temporal annotations, grouped by temporal category and ordered by impact (IMP).

Appendix D. Requirements

D.1. Requirement sets

RID	Temporal Annotations
R1	A1, A2, A9, A10, A11, A20, A23, A29, A30, A31, A32, A35, A38, A40, A43, A47, A54, A55, A57, A58, A60, A63, A68
R2	A2, A63
R3	A2, A26, A30, A31, A32, A33, A68
R4	A2, A4, A5, A6, A8, A21, A26, A32, A45, A46, A68
R5	A2
R6	A2, A9, A23, A32, A38, A40, A41, A42, A43, A44, A45, A54, A58, A60, A63, A68
R7	A2, A14, A32, A38, A39, A43, A51, A52, A57, A59, A61, A62, A63, A64, A68
R8	A2, A11, A23, A32, A38, A39, A57, A59, A63, A64, A68
R9	A2, A3, A16, A17, A26, A30, A31, A32, A33, A35, A36, A40, A58, A63, A68
R10	A2, A4, A5, A6, A7, A8, A9, A11, A18, A20, A21, A22, A24, A25, A26, A27, A28, A29, A30, A31, A32, A33, A34, A35, A36, A37, A38, A41, A42, A43, A47, A48, A49, A51, A52, A53, A58, A66, A68
R11	A2, A38, A41, A43, A50, A51, A58, A60, A63, A68
R12	A2, A12, A13, A23, A32, A38, A68
R13	A2, A11, A14, A23, A32, A38, A39, A52, A57, A59, A61, A62, A63, A64, A68
R14	A2, A11, A38, A39, A57, A59, A61, A63, A68
R15	A2, A5, A7, A21, A26, A31, A32, A33, A36, A51, A68
R16	A2, A32, A51, A63
R17	A2, A41, A43, A50, A51, A58, A60, A63, A68
R18	A2, A11, A14, A23, A31, A32, A38, A39, A40, A43, A51, A52, A57, A59, A61, A62, A63, A64, A68
R19	A2, A4, A5, A6, A7, A8, A18, A20, A21, A22, A24, A25, A26, A27, A28, A29, A31, A32, A38, A40, A41, A43, A47, A48, A51, A52, A53, A57, A59, A61, A64, A68
R20	A6, A22, A35, A40, A42, A60, A63, A65, A68
R21	A9, A38, A41, A54, A55, A57, A68
R22	A9, A14, A38, A39, A40, A54, A55, A59, A61, A63, A64, A68
R23	A9, A23, A32, A38, A39, A54, A55, A57, A59, A61, A63, A64, A68
R24	A10, A26, A29, A31, A32, A35, A38, A39, A40, A57, A59, A62, A64, A68
R25	A11, A38, A39, A41, A43, A44, A59, A68
R26	A11, A20, A32, A35, A38, A40, A42, A54, A59, A63, A64, A68
R27	A11, A14, A18, A19, A23, A26, A32, A35, A38, A39, A51, A57, A59, A61, A62, A63, A64, A68
R28	A11, A26, A32, A41, A43, A57, A68
R29	A11, A17, A19, A20, A26, A32, A33, A35, A39, A68
R30	A11, A39, A43, A57, A59, A61, A62, A63, A64, A68
R31	A11, A31, A32, A38, A57, A68
R32	A11, A38, A39, A42, A43, A57, A59, A61, A63, A64, A68
R33	A11, A31, A38, A40, A41, A43, A52, A57, A59, A60, A61, A62, A63, A67, A68
R34	A11, A23, A41, A57
R35	A14, A23, A32, A38, A39, A41, A43, A57, A59, A61, A62, A63, A64, A68
R36	A15, A66, A68
R37	A19, A26, A43, A45, A46, A68
R38	A20, A22, A26, A32, A35, A54, A68
R39	A23, A32, A38, A43, A57, A58, A68
R40	A23, A32, A41, A43, A68
R41	A23, A32, A38, A39, A41, A57, A59, A61, A62, A63, A64, A68
R42	A23, A32
R43	A23, A26, A68
R44	A23, A32, A38, A59, A61, A63, A64, A68
R45	A23, A38, A39, A57, A59, A63, A68
R46	A26, A68
R47	A26, A29, A31, A32, A35, A66, A68
R48	A26, A31, A32, A68
R49	A31, A68
R50	A32, A39, A68
R51	A32, A68
R52	A32, A38, A41, A59
R53	A32, A58, A68
R54	A38, A41, A57
R55	A38, A42, A58, A60, A63, A68
R56	A38, A57, A63, A64
R57	A38, A59
R58	A38, A68
R59	A38, A39, A60, A68
R60	A38, A40, A43, A60, A63, A65, A68
R61	A39, A57, A61, A68

RID	Temporal Annotations
R62	A40, A63, A68
R63	A40, A59, A64
R64	A40, A42, A44, A45, A58, A68
R65	A43, A51, A68
R66	A43
R67	A51, A68
R68	A51
R69	A54
R70	A54, A55, A57, A68
R71	A56, A57, A59, A68
R72	A57, A59
R73	A60, A68
R74	A63, A68
R75	A68

Table 19: The full list of requirement sets considered by our survey. A language is defined as a set of annotations.

D.2. Requirements Importance

R	ON	PON	OC	POC	MAI	IMP
R19	1	0.01	40	0.41	0.53	0.60
R18	1	0.01	49	0.50	0.36	0.56
R10	1	0.01	39	0.40	0.46	0.55
R1	1	0.01	34	0.35	0.47	0.52
R75	84	0.86	17	0.17	0.12	0.51
R27	1	0.01	38	0.39	0.35	0.48
R6	1	0.01	32	0.33	0.30	0.40
R7	2	0.02	36	0.37	0.24	0.40
R9	1	0.01	32	0.33	0.25	0.37
R35	1	0.01	34	0.35	0.19	0.34
R33	1	0.01	28	0.29	0.24	0.33
R13	2	0.02	33	0.34	0.18	0.33
R74	26	0.27	18	0.18	0.17	0.32
R15	2	0.02	31	0.32	0.14	0.29
R58	30	0.31	18	0.18	0.11	0.29
R65	8	0.08	24	0.24	0.18	0.29
R8	3	0.03	29	0.30	0.14	0.28
R11	1	0.01	31	0.32	0.13	0.28
R24	1	0.01	31	0.32	0.13	0.28
R23	1	0.01	31	0.32	0.13	0.28
R17	2	0.02	29	0.30	0.14	0.28
R51	31	0.32	18	0.18	0.07	0.27
R41	2	0.02	31	0.32	0.11	0.27
R73	10	0.10	18	0.18	0.17	0.26
R67	13	0.13	21	0.21	0.12	0.26
R3	3	0.03	27	0.28	0.11	0.25
R14	3	0.03	25	0.26	0.13	0.25
R46	19	0.19	21	0.21	0.07	0.25
R44	7	0.07	23	0.23	0.12	0.25
R26	1	0.01	25	0.26	0.14	0.24
R47	2	0.02	26	0.27	0.12	0.24
R45	8	0.08	22	0.22	0.12	0.24
R62	10	0.10	19	0.19	0.13	0.24
R22	1	0.01	24	0.24	0.13	0.23
R38	1	0.01	24	0.24	0.12	0.23
R61	11	0.11	18	0.18	0.12	0.22
R70	4	0.04	19	0.19	0.15	0.22
R55	2	0.02	21	0.21	0.14	0.22
R4	1	0.01	24	0.24	0.11	0.22
R28	1	0.01	24	0.24	0.11	0.22
R32	1	0.01	25	0.26	0.10	0.22
R60	1	0.01	23	0.23	0.12	0.22
R12	1	0.01	22	0.22	0.13	0.22
R59	1	0.01	20	0.20	0.14	0.21
R37	1	0.01	23	0.23	0.11	0.21
R49	14	0.14	19	0.19	0.07	0.21
R71	1	0.01	19	0.19	0.15	0.21
R48	8	0.08	25	0.26	0.04	0.21
R29	1	0.01	24	0.24	0.10	0.21
R39	2	0.02	23	0.23	0.10	0.21
R21	1	0.01	22	0.22	0.11	0.21
R25	1	0.01	21	0.21	0.12	0.20
R30	2	0.02	22	0.22	0.10	0.20
R20	1	0.01	21	0.21	0.11	0.20
R64	2	0.02	18	0.18	0.13	0.20
R50	11	0.11	19	0.19	0.05	0.19
R31	3	0.03	22	0.22	0.07	0.19
R40	3	0.03	21	0.21	0.07	0.18
R43	2	0.02	22	0.22	0.07	0.18
R53	6	0.06	19	0.19	0.07	0.18
R5	19	0.19	1	0.01	0.13	0.14
R36	1	0.01	18	0.18	0.05	0.14
R16	3	0.03	4	0.04	0.18	0.14
R57	20	0.20	1	0.01	0.11	0.13
R72	17	0.17	1	0.01	0.12	0.13
R56	10	0.10	1	0.01	0.14	0.12
R54	6	0.06	1	0.01	0.16	0.11
R2	12	0.12	2	0.02	0.11	0.11
R63	6	0.06	1	0.01	0.13	0.09
R66	21	0.21	1	0.01	0.02	0.08

R	ON	PON	OC	POC	MAI	IMP
R68	15	0.15	1	0.01	0.02	0.06
R42	14	0.14	1	0.01	0.02	0.05
R52	4	0.04	2	0.02	0.06	0.05
R34	1	0.01	1	0.01	0.08	0.04
R69	9	0.09	1	0.01	0.02	0.03

Table 20: The full list of requirements ordered by the their importance (IMP). ON: Number of ontologies for which requirement set is necessary. PON: ON as proportion. OC: Number of ontologies which are completely covered by requirement set. POC: OC as proportion. MAI: Mean importance of annotations in requirement set. IMP: Overall importance of requirement set. Shaded in gray or those requirements which are on the Pareto frontier wrt to PON, POC and IMA.