Aligning Conference Ontologies with SUMO: A Report on Manual Alignment via WordNet

Daniela SCHMIDT [a], Adam PEASE [b], Cassia TROJAHN [c] and Renata VIEIRA [a]

[a] *Pontifical Catholic University of Rio Grande do Sul (Brazil)*

[b] *Infosys, Foothill Research Center (USA)*

[c] *Institut de Recherche en Informatique de Toulouse (France)*

Abstract. This paper presents the process of manually establishing alignments between domain and foundational ontologies. The ontologies from the OAEI Conference track have been aligned to the SUMO foundational ontology. The Conference dataset is one of the most used dataset in ontology matching evaluation and has been extended in several versions. However, it lacks in alignments to foundational ontologies. As a complete manual alignment between SUMO and WordNet is available, we use such alignments as bridges to facilitate the matching task. In this paper we describe the constructing of such alignments and discuss the issues dealt with during the process, lessons learned and perspectives in the field.

Keywords. foundational ontologies, ontology matching, WordNet, SUMO

1. Introduction

Foundational ontologies play a fundamental role in the construction and integration of domain ontologies, providing a reference model that can be shared across domains. The clarity in semantics and the rich formalization of foundational ontologies are important requirements for ontology development in general [17,13], since it improves ontology quality. They also act as semantic bridges supporting interoperability between domain ontologies [15,13,18]. There are two approaches for the use of foundational ontologies [28]. With a *top-down approach*, the foundational ontology is used as a reference for deriving domain concepts, taking advantage of the knowledge and experience already encoded in it. In a *bottom-up approach*, one usually matches an existing domain ontology to the foundational ontology. The latter is more challenging since inconsistencies may exist between domain and foundational ontologies and one has to deal with different levels of abstraction in the matching process. This paper focus on the latter.

This paper discusses the alignment of the Conference ontologies [35] to a foundational ontology. The Conference dataset has been used in the Ontology Alignment Evaluation Initiative (OAEI)[1], which have been carried out over the last fifteen years. The OAEI, however, still lacks matching tasks involving foundational ontologies. The choice

[1] `http://oaei.ontologymatching.org/`

for aligning this OAEI dataset to a foundational ontology is motivated by the fact that it has became one of the most used in matching evaluations [35]. This dataset has also been extended with different proposals [4,16], and recently it is also covering complex alignments [33].

We present the process of establishing a consensual alignment between the Conference ontologies and SUMO (*Suggested Upper Merged Ontology*) [19,21]. Matching foundational and domain ontologies is far from being a trivial task and most approaches still rely on manually or semi-automatic strategies. This has been corroborated in [31], where manually classifying domain entities under foundational ontology concepts is reported to be very difficult to do correctly. As knowledge on foundational ontologies is highly specialized, it is important that such alignments consider the participation of different experts in the area. The findings in [31] also point out the need for improving the methodological process of manual integration of domain and foundational ontologies, in accord with what has been stated in [13]. As a complete manual alignment between SUMO and WordNet has been previously provided [20] and continually updated since the original effort, we argue here that using these alignments as bridges to matching domain ontologies to SUMO can facilitate the matching task.

We have chosen SUMO for several reasons. It is the only formal ontology that has a complete set of manually-performed correspondences to all 117,000 word senses in WordNet. It is also one of the few ontologies that has a detailed formalization in an expressive logical language. Most ontologies are still simple taxonomies and frame systems, and so assessing the meaning of their terms requires human intuition based on term names and relationships. SUMO includes a computational toolset [23] that allows users to test the logical consistency of its definitions, which provides a guarantee of quality and correctness than just testing type constraints. Lastly, SUMO is large and comprehensive at roughly 20,000 terms and 80,000 hand-written logical axioms, exceeding the size of other open source foundational ontologies by several orders of magnitude.

In this paper we describe the design choices and methodology followed for constructing this bridged alignment. We discuss the main issues that the experts were faced with during the process and the lessons learned and perspectives in the field. The rest of this paper is organised as follows. Section 2 discusses the main related works on matching domain and foundational level ontologies and the manual construction of alignments. Section 3 describes the domain and foundational ontologies used here, together with the alignments between SUMO and WordNet. Section 4 presents the methodology we followed to match the domain ontologies to WordNet. Section 5 discusses the main issues we have faced with and finally Section 6 concludes the paper and discusses future directions.

2. Related Work

Matching Domain Ontologies to Foundational Ontologies As reported in [14,13], methodologies for constructing ontologies should not neglect the use of foundational ontologies and should better address it in a *top-down* approach. In the absence of systematic adoption of foundational ontologies within the domain ontology development process, *bottom-up* approaches have to be applied instead. Matching domain and foundational ontologies has been done mostly manually. In [3], geoscience ontologies (GeoSciML and

SWEET – Semantic Web for Earth and Environmental Terminology) have been manually aligned to DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*) [8][2] and incompatibilities issues have been discussed. In [17], DOLCE has also been manually aligned to a domain ontology describing services (OWL-S), in order to address its conceptual ambiguity, poor axiomatization, loose design and narrow scope. In [5], several schemata of FactForge (which enables SPARQL querying over the Linked Open Data cloud) have been aligned to PROTON (PROTo ONtology) [32][3] in order to provide a unified way to access the data. The alignments were created by knowledge engineers, where equivalence and subclass relationships between DBPedia, Geonames and Freebase were established to PROTON classes. They report the different strategies adopted and how the ontologies had fit them (for instance, the fact that reference instances were not included in the FactForge datasets and the need of using alternative strategies or extending the dataset with the necessary instances). Contrary to us, they have modified the original ontoloiges. Manually alignments have also been established between biomedical ontologies and BFO (*Basic Formal Ontology*)[4] [10,1], in [29]. Although many proposals have focused on DOLCE, it lacks complete alignments to WordNet, as found in SUMO, also it is several orders of magnitude smaller. PROTON and BFO are also limited, they present formal axioms in an expressive logic but are small in size.

While those proposals mainly generate manual alignments between foundational and domain ontologies, one of the few automatic approaches is BLOOMS+ [12]. It has been used to automatically align PROTON to LOD datasets using as gold standard the alignments provided in [5]. BLOOMS+ first uses Wikipedia to construct a set of category hierarchy trees for each class in the source and target ontologies. It then determines which classes to align using 1) similarity between classes based on their category hierarchy trees; and 2) contextual similarity between these classes to support (or reject) an alignment. BLOOMS+ significantly outperformed existing matchers in the task.

More recently, an automatic approach for matching domain and foundational ontologies has been proposed in [26]. It exploits existing alignments between WordNet and foundational ontologies. The matching process is divided in two main steps. The first step identifies the correct synset to the domain concept and the second one identifies the correspondence the domain concept to a foundational concept. The approach has been evaluated using DOLCE and domain ontologies from the OAEI conference data set[5], with the help of the alignments provided in [9,20]. This work has been further extended in [27], where two similarity measures for synset disambiguation have been adopted: (1) an adaptation of the Lesk measure and (2) word embeddings. The evaluation has been also extended including DOLCE and SUMO ontologies and their alignments to WordNet and three domain ontologies (SSN – Semantic Sensor Network ontology Core Ontology for Robotics and Automation, and OAEI Conference). Here, we discuss the consensual process of manually constructing the alignments between SUMO and OAEI Conference dataset.

Consensual Alignments While different ontology alignments have been constructed from manual analysis, involving a different number of experts and resulting in different

[2] http://www.loa.istc.cnr.it/old/DOLCE.html
[3] http://ontotext.com/proton
[4] https://github.com/bfo-ontology/BFO/wiki
[5] http://oaei.ontologymatching.org/2017/conference/index.html

levels of agreement, the focus has mostly been on describing the resulting alignment rather than on the details of the manual process. Guidelines for constructing alignments are in fact scarce in the field, though there are more general discussions on the qualities of a good benchmark in other research fields [30,6]. It may be obvious that the fundamental problem of aligning ontologies is determining what is the meaning of the terms that are candidates for alignment. If the meaning is implicit, and one must resort to the domain knowledge of human matchers, then only an automatic suggestion is feasible. This is even more required when dealing with foundational ontologies.

Construction of alignments in general follows different strategies, including starting the alignment generation from scratch, relying on a set of initial alignments for gathering additional ones, and creating a reference from validating and selecting a set of correspondences from automatically generated correspondences from a number of matching systems. In the first category, the creation of the first reference alignment of the Conference dataset dates back to 2008, when the track organizers created a reference alignment for all possible pairs of five of the conference ontologies. The reference alignments were based on the majority opinion of three evaluators and were discussed during a consensus workshop. This dataset has evolved over the years, as described in [35], with the feedback from the OAEI participants and has been revised in [4]. They re-examined the dataset with a focus on the degree of agreement between the reference alignments and the opinion of experts. With the aim of studying the way different raters evaluate correspondences, in [34] experiments in manual evaluation have been carried out using a set of correspondences generated by different matchers between different vocabularies. Five raters evaluated alignments and talked through their decisions using the think aloud method. Their analysis showed which variables can be controlled to affect the level of agreement, including the correspondence relations, the evaluation guidelines and the background of the raters. That work refers as well to the different levels of agreements between annotators reported in the literature. While a perfect agreement between raters is reported in the Very Large Crosslingual dataset in [7], [11] reported a quite different observation when establishing *owl:sameAs* relationships in the LOD. These aspects have also been discussed in [31] for the task of integrating foundational and domain ontologies.

3. Ontologies and WordNet Alignments

3.1. Foundational and Domain Ontologies

3.1.1. SUMO

SUMO [19][6][7] (*Suggested Upper Merged Ontology*) is a comprehensive ontology that includes general concepts and their definitions as well as a number of integrated domain ontologies. It totals approximately 20,000 terms and 80,000 logical axioms. It is defined in a higher-order logic, which supports mathematical statements about modality, belief, likelihood and temporal qualification. It is this expressiveness of the logic used that provides unambiguous formal definitions of terms, without recourse to human interpretation. SUMO includes an associated tool set called Sigma [23] that is used to test

[6] `http://www.ontologyportal.org`
[7] `https://github.com/ontologyportal/sumo`

Table 1. Classification of OAEI Conference ontologies.

Ontology	Type	#Concepts	#Top concepts	DL expressivity
Cmt	Tool	36	8	$ALCIN(D)$
ConfTool	Tool	38	7	$SIN(D)$
Edas	Tool	104	16	$ALCOIN(D)$
Ekaw	Insider	74	6	$SHIN$
Iasted	Web	140	10	$ALCIN(D)$
Sigkdd	Web	49	9	$ALEI(D)$
Sofsem	Insider	60	14	$ALCHIF(D)$

and extend the ontology and deploy it for applications involving automated inference. It includes automatic translators to the following languages: TPTP (for first order logic with equality) [24,25], TFF0 (for first order logic with arithmetic) [22] and THF [2] (for classical higher-order logic) used by major classes of automatic theorem provers.

3.1.2. OAEI Conference Dataset

The Conference dataset[8] has been used to evaluate nearly all matching systems developed [4] and it is quite a challenging dataset in the field [35]. This dataset is composed of 16 ontologies on the conference organization domain and simple reference alignments between 7 of these ontologies. They cover different aspects of the domain and are classified in three categories: i) *Web* (conference series and its web pages), ii) *Tool* (software tool for conference organisation support), and iii) *Insider* (experience of people with personal participation in conference organisation). Here, we used the 7 ontologies for which the reference alignments are available. Table 1 presents the classification of the ontologies and their number of concepts. Although they differ in DL expressivity, this was not taken into account in our manual alignment methodology.

3.2. SUMO to WordNet Alignment

In alignment between SUMO and WordNet, for each identified correspondence, the synset of WordNet is augmented with three pieces of information: *(i)* a prefix (&%) that indicates that the term is taken from SUMO; *(ii)* the SUMO concept; and *(iii)* a suffix indicating the kind of relation. The suffix '=' indicates that the correspondence relation is synonymy. '+' indicates that the concept is a hypernym of the associated synset. The instantiation relation is indicated by the suffix '@'. An example of the structure of a correspondence representing a synonymy relation can be seen below. In the example, *"02761392 06 n 03 automaton 0 robot 0 golem"* corresponds to the synset. The gloss is defined as *"a mechanism that can move automatically"*, the prefix *"&%"* indicates that the term is taken from SUMO. *"Device"* corresponds to the SUMO concept and the signal *"+"* is the suffix indicating the hyponymy relation.

02761392 06 n 03 automaton 0 robot 0 golem a mechanism that can move automatically &%Device+

[8] `http://oaei.ontologymatching.org/2016/conference/index.html`

4. Alignment Methodology

This section describes the overall methodology we have followed to create the consensual alignment between SUMO and the domain ontologies. As stated above, this process relies on the existing alignment between SUMO and WordNet. We consider these previous alignments as correct so the aim here is to identify the right WordNet synset to the domain concepts. Also, we have reduced the problem to the first-level concepts of the hierarchies from the domain ontologies. This has resulted in 70 first-level domain concepts (Table 1). For each first level concept of the domain ontology, a foundational specific concept is associated. The cost of doing manual alignment with first level concepts is smaller, as it is reduced to the number of concepts at the first level.

Four evaluators have been involved in the task of aligning the 70 top-level domain concepts to the WordNet synsets. The evaluators are researchers, therefore all have common-sense knowledge about conferences (the domain ontology), they have background in Computer Science and are well acquainted with ontology matching. One of the evaluators is the creator of the SUMO ontology.

The overall methodology is articulated in the following two steps: i) Individually generating the alignments between domain concepts and WordNet synsets; ii) Collaboratively validating the set of found correspondences. Next we detail each step.

4.1. Individual Generation of Correspondences

In this first step, each evaluator aligned each of 70 domain concepts to WordNet synsets. To that extent, each domain concept and the corresponding WordNet synsets, resulting from searching in WordNet for the term associated to the domain concept, were listed to the evaluators. In the absence of entries in WordNet for the terms, a head modifier strategy has been applied (i.e., WrittenPaper is a Paper). Only one concept had not corresponding entry in WordNet (*sigkdd#Sponzor*[9]). In order to help the evaluator to understand the context of the domain concept, their sub-concepts were also presented. As the domain ontologies are equipped with very few comments or labels, we have completed the description of the concept using the definitions from the Cambridge Dictionary[10]. However, we are aware that the found definitions may not reflect the exact semantic of the concept. Each evaluator then was asked to select the right WordNet synset for each domain concept. The evaluators were instructed to select one option for each domain concept, however, in some cases more than one sense was selected. This happens because the domain concept was not clear enough, or the senses available in WordNet were very general. Evaluators were also invited to comment their decisions. Table 2 shows a fragment of the spreadsheet for the domain concept *cmt#Conference*.

4.2. Validating the Correspondences

After the individual annotation of each domain concept with the WordNet synset, the annotators were able to see the annotations of each other and identify the conflicts. Based on the views on the other annotators (and their comments), each one was able to change their initial annotations. For those conflicts where the comments were not enough for

[9] `http://oaei.ontologymatching.org/2018/conference/data/sigkdd.owl`
[10] `http://dictionary.cambridge.org/us/`

Table 2. Example of spreadsheet adopted by the evaluators.

Domain concept	WordNet synsets	SUMO concept	Comments
cmt#Conference	**S1**: conference group discussion a discussion among participants who have an agreed (serious) topic **S2**: league conference an association of sports teams that organizes matches for its members **S3**: conference a prearranged meeting for consultation or exchange of information or discussion (especially one with a formal agenda)	1 – Communication+ 2 – SportsLeague+ 3 – FormalMeeting=	An academic conference would be a FormalMeeting

Domain concept	WordNet synsets	SUMO concept
cmt#Decision	S1: decision determination conclusion S2: decision determination conclusion S3: decisiveness decision S4: decision S5: decision	1: **Learning+** 2: **Deciding+** 3: TraitAttribute+ 4: ConstantQuantity+ 5: ConstantQuantity+
cmt#Preference	S1: predilection preference orientation S2: preference S3: preference predilection taste S4: preference druthers	1: **IntentionalRelation+** 2: SubjectiveAssessmentAttribute+ 3: PsychologicalAttribute+ 4: **PsychologicalAttribute+**
edas#PersonalHistory	S1: history S2: history S3: history account chronicle story S4: history S5: history	1: **PastFn=** 2: History= 3: **HistoricalAccount+** 4: PastFn= 5: Proposition+
sigkdd#Award	S1: award awarding S2: award accolade honor honour laurels S3: prize award S2: function office part role	1: UnilateralGiving+ 2: **ContentBearingObject+** 3: **UnilateralGiving+** 2: Position+
Conference#Review_preference	S1: predilection preference orientation S2: preference S3: preference penchant predilection taste S4: preference druthers	1: **IntentionalRelation+** 2: SubjectiveAssessmentAttribute+ 3: PsychologicalAttribute+ 4: **PsychologicalAttribute+**

Table 3. Conflicts solved after discussion between annotators. Boldface concepts represent the conflicts and final results are indicated with a underline.

understanding the annotation, an online discussion took place. From the 70 annotated domain concepts, 8 of them have been annotated with different WordNet synsets. Table 3 lists some examples, which are discussed in the next section. All generated alignments are available in the Alignment API format[11].

5. Discussion

During the process of alignment construction, several difficulties arose for interpreting the real meaning that the concept represents in the domain ontology. For instance, the concepts `Bid` and `Preference` (Table 3) in `cmt` ontology had no description clarifying its use, and no sub or super concepts which could be used to clarify their meaning. In these cases, the evaluators discussed and considered the proper meaning according to their own interpretation of the domain, however, such cases may interfere with the quality of the resulting reference alignment because there is no objective standard for what the meaning, and therefore the correct mapping must be. We have only the consensus guess about intended meaning among human evaluators. In addition, some concepts rep-

[11] https://github.com/danielasch/ReferenceAlignment

resented in the ontology present other kind of problems such as doubts regarding ontology elements' adequacy, for example, the concept ReviewRating in edas ontology, which according to the discussion raised by the evaluators, a rating could be a relationship between a thing, an agent and a rating value. In the same way, the concept Deadline in sigkdd ontology could be a relationship between the conference and a date. They are however defined in those ontologies as concepts, rather than relationships.

In other cases, sub-concepts are different from first-level concepts and therefore they represent different information, as the concept Event in ConfOf ontology. Some of their sub-concepts Social_event/Banquet, Working_event/Conference, Working_event/Workshop are in line with the main concept, however others such as Administrative_event/Camera_Ready_event seems out of the context. In fact, it should not be a Process at all but a deadline for doing something (submitting a version of a paper, for instance).

In contrast, one can examine a SUMO definition of a term such as FormalMeeting[12] and see that it is necessarily a Meeting that is not a SocialParty, that it must be temporally preceded by a Planning that has the result of creating the meeting, as well as constraints that other events like a Resolution to be considered such, may only occur at a FormalMeeting. Something like a modern dictionary, but with the definitions expressed in logic, rather than human language, so that a machine can perform computation (and consistency checking) with those definitions. The cases described above consist of ontological representation problems commonly present in lightweight ontologies, and hinder the reuse and reliability of the represented knowledge. In addition, they highlight the importance of advancing in research that uses top-level ontologies to give more formalization to domain ontologies.

The challenges in aligning the OAEI ontologies should highlight two elements that are lacking in the majority of most of current ontology practice. The first element is the degree of reuse. Ontologies that are created from scratch suffer from the fact that their terms have only a small number of relationships to other terms. The point of having an ontology is to have a shared meaning among its users. When domain ontologies are created in isolation, rather than as extensions to widely used comprehensive ontologies they miss an opportunity for sharing common meaning. Modern software development, for example in Java or Python, means reusing vast amounts of existing code, such as extensive language libraries and other packages like web servers, databases, device drivers etc. Ontology development needs to follow the same practice to achieve the same efficiency of process as procedural software development.

A second element is the expressivity of definitions. If, as with several of the OAEI ontologies, one must guess at the intended meaning of a term only by its name, then there isn't much chance for shared meaning amongst its users. Each user will just be making a guess. If each term has only a set of binary relationships to other terms then it should still be clear that issues like mutual constraints on values and boundary cases are left unformalized and also at risk of being in conflict among its users. Additionally, without a computational formalization of such constraints, the computer will not be able to test or enforce them. Comments in natural language, no matter how extensive or precise, will not overcome the need for computational definitions, and our experience in this matching effort has been that comments are often not even present, and rarely extensive or precise.

[12] http://sigma.ontologyportal.org:8080/sigma/Browse.jsp?lang=EnglishLanguage&flang=SUO-KIF&kb=SUMO&term=FormalMeeting

In our work, we were not engaged in correcting the ontologies, since they are part of public datasets. However we consider that a discussion about the problems identified is necessary. Perhaps more robust alignment processes would inherently require modifications in target domain ontologies but also certainly a more detailed formalization. Given the paucity of definitions, we are limited primarily to linguistically-based matches and use of WordNet is a suitable choice for assisting with this sort of match.

6. Conclusions

This paper has discussed the alignment between domain ontologies from the Conference domain and SUMO. One of the main issues experts have been faced concerns the lack of formal definitions associated to rich terminological layers (comments and labels) helping to understand the precise semantics of each concept. We claim that domain ontologies should rather be developed on the top of such ontologies. This work is a first step toward the construction of a dataset involving foundational ontologies that can serve to evaluation systems in the context of OAEI campaigns. These alignments can also be explored as semantic bridges in domain ontology matching.

References

[1] R. Arp, B. Smith, and A. Spear. *Building Ontologies with Basic Formal Ontology*. MIT Press, 2015.

[2] Christoph Benzmüller and Adam Pease. Progress in Automating higher-order ontology reasoning. In Boris Konev, Renate Schmidt, and Stephan Schulz, editors, *Proceedings of the Workshop on Practical Aspects of Automated Reasoning*. Edinburgh, UK, 2010.

[3] B. Brodaric and F. Probst. DOLCE ROCKS: Integrating Geoscience Ontologies with DOLCE. In *Semantic Scientific Knowledge Integration*, pages 3–8, 2008.

[4] Michelle Cheatham and Pascal Hitzler. Conference v2. 0: An uncertain version of the oaei conference benchmark. In *International Semantic Web Conference*, pages 33–48. Springer, 2014.

[5] M. Damova, A. Kiryakov, K. Ivanov Simov, and S. Petrov. Mapping the central LOD ontologies to PROTON upper-level ontology. In *Workshop on Ontology Matching*, 2010.

[6] A Dekhtyar and J Huffman Hayes. Good benchmarks are hard to find: Toward the benchmark for information retrieval applications in software engineering. In *Proceedings of the 22th International Conference on Software Maintenance*, 2006.

[7] Jérôme Euzenat, Alfio Ferrara, Laura Hollink, Antoine Isaac, Cliff Joslyn, Véronique Malaisé, Christian Meilicke, Andriy Nikolov, Juan Pane, Marta Sabou, François Scharffe, Pavel Shvaiko, Vassilis Spiliopoulos, Heiner Stuckenschmidt, Ondrej Sváb-Zamazal, Vojtech Svátek, Cássia Trojahn dos Santos, George A. Vouros, and Shenghui Wang. Results of the ontology alignment evaluation initiative 2009. In *Proceedings of the 4th International Workshop on Ontology Matching*, 2009.

[8] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening Ontologies with DOLCE. In *13th Conference on Knowledge Engineering and Knowledge Management*, pages 166–181, 2002.

[9] A. Gangemi, R. Navigli, and P. Velardi. *The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet*, pages 820–838. 2003.

[10] P. Grenon, B. Smith, and L. Goldberg. Biodynamic Ontology: Applying BFO in the Biomedical Domain. In *Stud. Health Technol. Inform*, pages 20–38. IOS Press, 2004.

[11] Harry Halpin, Patrick J. Hayes, James P. McCusker, Deborah L. McGuinness, and Henry S. Thompson. When owl:sameas isn't the same: An analysis of identity in linked data. In *Proceedings of the 9th International Semantic Web Conference*, pages 305–320, 2010.

[12] P. Jain, P. Yeh, K. Verma, R. Vasquez, M. Damova, P. Hitzler, and A. Sheth. Contextual Ontology Alignment of LOD with an Upper Ontology: A Case Study with PROTON. In *Proceedings of the 8th Extended Semantic Web Conference*, pages 80–92, 2011.

[13] C. Keet. The use of foundational ontologies in ontology development: An empirical assessment. In *Proceedings of the 8th Extended Semantic Web Conference*, pages 321–335, 2011.

[14] Z. Khan and C. Keet. ONSET: Automated Foundational Ontology Selection and Explanation. In *Proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management*, pages 237–251, 2012.

[15] V. Mascardi, A. Locoro, and P. Rosso. Automatic Ontology Matching via Upper Ontologies: A Systematic Evaluation. *IEEE Trans. on Knowl. and Data Eng.*, 22(5):609–623, 2010.

[16] Christian Meilicke, Raul Garcia-Castro, Fred Freitas, Willem Robert van Hage, Elena Montiel-Ponsoda, Ryan Ribeiro de Azevedo, Heiner Stuckenschmidt, Ondrej Sváb-Zamazal, Vojtech Svátek, Andrei Tamilin, Cássia Trojahn, and Shenghui Wang. Multifarm: A benchmark for multilingual matching. *JWS*, 15:62–68, 2012.

[17] P. Mika, D. Oberle, A. Gangemi, and M. Sabou. Foundations for Service Ontologies: Aligning OWL-S to DOLCE. In *Proceedings of the 13th Conference on World Wide Web*, pages 563–572, 2004.

[18] Julio Cesar Nardi, Ricardo de Almeida Falbo, and João Paulo A. Almeida. Foundational ontologies for semantic integration in EAI: A systematic literature review. In *Proceedings of the 12th IFIP WG Conference on e-Business, e-Services, and e-Society, I3E*, pages 238–249, 2013.

[19] I. Niles and A. Pease. Towards a Standard Upper Ontology. In *Conference on Formal Ontology in Information Systems*, pages 2–9, 2001.

[20] I. Niles and A. Pease. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In *Proceedings of the Inter. Conference on Knowledge Engineering*, pages 412–416, 2003.

[21] Adam Pease. *Ontology: A Practical Guide*. Articulate Software Press, Angwin, CA, 2011.

[22] Adam Pease. Arithmetic and inference in a large theory (extended abstract). In *4th Conference on Artificial Intelligence and Theorem Proving*, 2019.

[23] Adam Pease and Christoph Benzmüller. Sigma: An Integrated Development Environment for Logical Theories. *AI Communications*, 26:9–97, 2013.

[24] Adam Pease and Stephan Schulz. Knowledge Engineering for Large Ontologies with Sigma KEE 3.0. In *The International Joint Conference on Automated Reasoning*, 2014.

[25] Adam Pease, Geoff Sutcliffe, Nick Siegel, and Steven Trac. Large Theory Reasoning with SUMO at CASC. *AI Communications, Special issue on Practical Aspects of Automated Reasoning*, 23(2-3):137–144, 2010.

[26] D. Schmidt, R. Basso, C. Trojahn, and R. Vieira. Matching Domain and Top-level Ontologies via OntoWordNet. In *Workshop on Ontology Matching*, pages 1–2, 2017.

[27] D. Schmidt, R. Basso, C. Trojahn, and R. Vieira. Matching domain and top-level ontologies exploring word sense disambiguation and word embedding. In *Emerging Topics in Semantic Technologies [best papers from the workshops at ISWC 2018]*, pages 27–38, 2018.

[28] S. Semy, M. Pulvermacher, and L. Obrst. Toward the use of an upper ontology for U.S. government and U.S. military domains: An evaluation. Technical report, MTR 04B0000063, The MITRE Corporation, 2004.

[29] V. Silva, M. Campos, J. Silva, and M. Cavalcanti. An Approach for the Alignment of Biomedical Ontologies based on Foundational Ontologies. *Information and Data Management*, 2(3):557–572, 2011.

[30] Susan Elliott Sim, Steve Easterbrook, and Richard C. Holt. Using benchmarking to advance research: A challenge to software engineering. In *Proceedings of the 25th International Conference on Software Engineering*, pages 74–83, 2003.

[31] R. Stevens, P. Lord, J. Malone, and N. Matentzoglu. Measuring expert performance at manually classifying domain entities under upper ontology classes. *Journal of Web Semantics*, 2018.

[32] I. Terziev, A. Kiryakov, and D. Manov. Base Upper-level Ontology (BULO) Guidance. Deliverable 1.8.1, sekt project, 2005.

[33] Élodie Thiéblin, Ollivier Haemmerlé, Nathalie Hernandez, and Cássia Trojahn. Task-oriented complex ontology alignment: Two alignment evaluation sets. In *Proceedings of the 15th Extended Semantic Web Conference*, pages 655–670, 2018.

[34] Anna Tordai, Jacco van Ossenbruggen, Guus Schreiber, and Bob Wielinga. Let's agree to disagree: On the evaluation of vocabulary alignment. In *Proceedings of the 6th International Conference on Knowledge Capture*, pages 65–72, 2011.

[35] Ondřej Zamazal and Vojtěch Svátek. The Ten-Year OntoFarm and its Fertilization within the Onto-Sphere. *Web Semantics: Science, Services and Agents on the World Wide Web*, 43:46–53, March 2017.