Towards a Semantic Matchmaking Algorithm for Capacity Exchange in Manufacturing Supply Chains

Audun Vennesland, Johannes Cornelis de Man, Peter Halland Haro, Emrah Arica and Manuel Oliveira
SINTEF, Trondheim, Norway

Keywords: Semantic Matching, Semantic Matchmaking, Manufacturing Supply Chains.

Abstract: Within supply chains, companies have difficulties in finding suppliers outside their known supplier pool or geographical areas. The EU project MANUSQUARE aims to deploy a marketplace to match supply and demand of supply chain resources to facilitate accurate and efficient matchmaking. To this end, a semantic matching algorithm has been developed as one of the key enablers of such a marketplace. The algorithm exploits formal descriptions of resources provided by an ontology developed in the project and will later be extended to incorporate additional data from different endpoints. This paper describes the main components of the semantic matching algorithm, which on the basis of the formally described supply chain resources returns a ranked list of relevant suppliers given a customer query. The paper further describes a comparative evaluation of a set of common semantic similarity techniques that was conducted in order to identify the most appropriate technique for our purpose. The results from the evaluation show that all four techniques perform pretty well and are able to distinguish relevant suppliers from irrelevant ones. The best performing technique is the edge-based technique Wu-Palmer.

1 INTRODUCTION

Complex and simple supply chains are continuously undergoing improvements in terms of efficiency and effectiveness, while exploring their integration with both new and existing supply chains in related areas. A key challenge within supply chain usage pertains to identifying and mapping the right supplier at the right time for the right service. Limited knowledge and trust concerns often make purchasers dependent on suppliers that are within reach: Local partners, well-known names within an industrial sector, Yellow-Pages or the first hits on a search engine. To establish supplier relationships outside a known domain or geographic area is therefore hard, creating limited value networks by utilizing suppliers that are limited to a specific geographic area and industrial domain. Although established supply chains are continuously optimized i.e. "the process of completing fulfillment functions efficiently and effectively" (Sherer, 2005), supply chains are not including suppliers that better match the required services to create better value networks, i.e. "link customer demand directly to their networked supply chains" (Sherer, 2005). A company could therefore benefit from matchmaking to establish supplier relationships outside its supplier pool or geographic area.

Accurate and efficient matchmaking of supply and demand of manufacturing resources, including physical assets as well as human know-how, can have a major economic impact on utilizing available capacity for the right product at the right time. A purchaser's job is made more efficient, using less time, and more effectively, finding the right supplier that delivers the right service at the right quality, and so on. Some of the more recent solutions have turned to semantic technologies for formal semantic descriptions that can be interpreted by machines to identify semantic similarity between offer and demand (Ameri and Patil, 2012; Järvenpää et al., 2018).

The EU project MANUSQUARE[1] aims to deploy a European Platform-enabled marketplace facilitating matchmaking of supply and demand of manufacturing resources. To support this objective, the project has developed the MANUSQUARE ontology for a formal representation of manufacturing resources. This ontology incorporates abstract concepts as well as domain-specific concepts capturing the knowledge of industrial sectors addressed in the project, such as manufacturing technologies solutions and textile and cosmetics production.

With the ultimate goal of optimising the match

[1] https://www.manusquare.eu/

Vennesland, A., Cornelis de Man, J., Haro, P., Arica, E. and Oliveira, M.
Towards a Semantic Matchmaking Algorithm for Capacity Exchange in Manufacturing Supply Chains.
DOI: 10.5220/0008364404660472
In *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2019)*, pages 466-472
ISBN: 978-989-758-382-7

between supply and demand of manufacturing resources, the semantic matchmaking algorithm considers two types of input:

1. Manufacturing resources offered by suppliers. These resources are represented as RDF triples in a knowledge base. The MANUSQUARE ontology enables a formal description of the resources.
2. Various data provided by different endpoints that can enhance the accuracy of the matchmaking and offer added-value functionality. This can include data from historical transactions to evaluate suppliers along different dimensions, e.g. historical matches, reputation indicators, consumer satisfaction, etc.

The focus of this paper is on the first type of input and the technical development described is a semantic matching algorithm that enables matching of offer and demand based their representation in a knowledge base according to the MANUSQUARE ontology.

The main contributions from this work are:

- An approach for a semantic matching algorithm that matches resource demand with available resources offered by suppliers.
- A summary of results from a comparative evaluation of different semantic matching techniques used in the algorithm.
- Ideas on how supplementary data from different sources could further enhance the results and contribute to innovation in manufacturing supply chains.

2 RELATED WORK

An initial literature study related to semantic matchmaking in manufacturing logistics gave limited results[2], however, we identified the three following papers as relevant. Ameri and Patil (Ameri and Patil, 2012) suggested a multi-agent framework based on formal semantics for connecting buyers and sellers of manufacturing services. In order to overcome the limitations of pure string equality-based similarity techniques, mediating agents match offered services with requested capabilities based on *taxonomy-based similarity* and *feature-based similarity*. The taxonomy-based similarity uses subsumption reasoning to determine if the ontological concept(s) representing a query can be addressed by ontological concept(s) representing available services. The feature-based similarity refines the results from the taxonomy-based similarity by utilising the logical constraints described by the concepts (e.g. via property restrictions) in order to more accurately rank the relevant services. A weighted set similarity measure, the Tversky Measure (Tversky, 1977), was used to compute a similarity score from which the offered services are ranked.

Järvenpää et al. (Järvenpää et al., 2018) developed a *capability matchmaking procedure* for matching product requirements with resource capabilities and possible combinations thereof. The matchmaking relied on a combination of ontologies allowing to formally express product requirements and resource capabilities, and business rules expressing more detailed parameters such as dimensions of a given resource. The matchmaking between product requirements and resource capabilities consists of two consecutive steps: (1) *Matching of product requirements and resource capabilities at concept level.* Product requirements are represented as individuals of concepts in the general capability ontology, and so are the available resources. Hence, there is a match between product requirement and available resources as long as they are members of the same ontology concept; (2) *Detailed matching of parameters.* This step employs the specified business rules and checks if there is a match between parameters specified in the product requirement and the offered resource (e.g., if a screw type used by a screwing machine resource complies with the required screw-type defined in the product requirement).

Schönböck et al. (Schönböck et al., 2018) used matchmaking in the context of volunteering, i.e., connecting tasks with volunteers willing and capable of performing them. An ontology coupled with meta-information enabling a more explicit definition of expertise or task preference was used as a basis for the matchmaking, resulting in a ranked list of tasks or volunteers whose profiles match as closely as possible. The ontology included core aspects such as competencies, spatio-temporal constraints and social relationships. Tasks and volunteers are represented as instances in the ontology and similarity values and meta-information are linked to properties in the ontology. The matchmaking score between a given task and a given volunteer is calculated based on (1) aggregating the similarity values associated with the relationships (properties) between concepts these instances are members of (explicit similarity), (2) the taxonomic structure of these concepts (implicit similarity) and (3) meta-information such as how much volunteers like/dislike a task, their level of expertise and how important the task is. The former is based on a fixed similarity value representing the strength

[2] The literature search was conducted in Elsevier Scopus and Google Scholar

of the relationship between two facet concepts (e.g. equivalentTo has a higher similarity value than relatedTo), whereas the latter is based on semantic similarity techniques considering e.g. the taxonomic proximity and depth of the concepts.

3 APPROACH

The semantic matching algorithm described in this section returns a ranked list of suppliers whose offered resources match a consumer query. Both the supplier resources and the consumer query are formally represented by a set of ontology concepts defined in the MANUSQUARE ontology. The ontology consists of over 1000 classes and a single resource record as well as a query can be represented in a multitude of ways. Figure 1 illustrates how a supplier resource can be formulated using these facets in the MANUSQUARE ontology. Here, the supplier Hackett Group, which is situated in Penedo, Portugal, offers welding of steel. This supplier can perform this process between 01.10.2019 and 01.12.2019 with an available capacity of 59 working hours. The company is certified according to the ISO9001 quality management standard.

Through discussion with academics and industry representatives within the project, the following facets are considered the most relevant parameters for the semantic matching:

- *Process (P):* which manufacturing process is sought by the customer?
- *Material (M):* which material does the consumer want to have processed?
- *Machine (MA):* which machine / equipment is required for performing the process?
- *Certifications (C):* which company and/or quality management certifications are required by the consumer?
- *Capacity (CP):* how much capacity (in terms of production hours) is required to perform the requested process?
- *Calendar Availability (CA):* Can the supplier deliver by the required due-date?

The facets are included as parameters in the following equation which returns a similarity score between a consumer query (q) and each individual resource registered in the knowledge base (r):

$$Sim(q,r) = \sum_{x \in \{F\}} sf(x) \tag{1}$$

F represents the set of facets values, and *sf* represents a similarity function as described in the following. The similarity score returned represents the average of the sub-scores computed for each facet, and when run on a dataset, the result is a ranked list of suppliers whose offered resources match the facets expressed in the query. Note that a consumer can search for a combination of different processes and leave blank uncertain fields in the query. This is considered when computing the aggregate similarity score.

For the facets Capacity and Calendar Availability the similarity between a query and available resources can be computed using simple Boolean matching, i.e. either the registered supplier resource meets the constraints expressed by the consumer query (score 1.0) or it does not (0.0).

For the facet Certifications, similarity is computed using the Jaccard set similarity measure (Jaccard, 1901). The Jaccard similarity is computed by finding the intersected set of certifications from a query and a supplier and dividing this by the union of certifications.

For the facets Process, Material, and Machine other similarity measures are needed. In order to compute a similarity score between the query and available resources along these facets, techniques able to exploit the taxonomic position as well as the context of a resource should be employed. Such techniques are typically categorised into edge-based and information content-based techniques (Jiang and Conrath, 1997). Edge-based techniques consider the path distance and taxonomic position of concepts to be matched, while information content-based methods are based on the assumption that the more information two concepts share, the more similar they are. Here, the information shared by two concepts is derived from the information content of the concept(s) that subsume them in the taxonomy. Information content is quantified as negative the log-likelihood of finding a given concept in a taxonomy and basically states that the more abstract a concept is, the less information it holds (Resnik, 1995). In our approach, we apply a variant of information content called *intrinsic information content* (Seco et al., 2004; Pirró and Talia, 2010) which contrasts the "conventional" information content by not relying on usage statistics of concepts in a corpus. The intrinsic information content of a concept c is computed as follows:

$$IC(c) = 1 - \frac{log(Sub(c) + 1}{log(|C|)} \tag{2}$$

where $Sub(c)$ indicates the number of subclasses of the concept c and $|C|$ represents the total number of

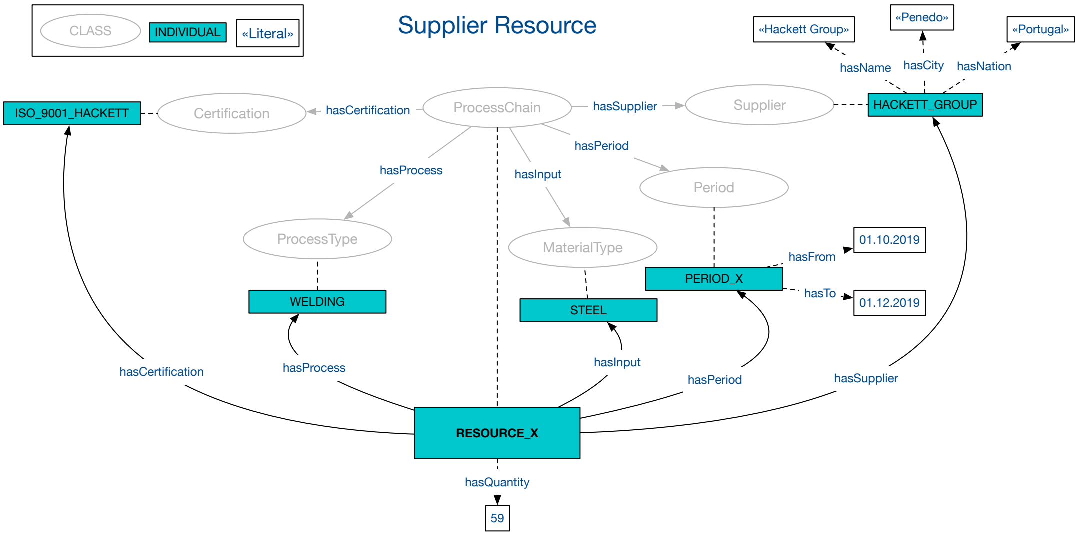


Figure 1: Ontological representation of a manufacturing resource.

concepts in the ontology.

The example illustrated in Figure 2 further explains the approach. For a given query, a similarity score is computed from a pairwise matching of the facets represented by the query and the facets represented by each resource stored in the knowledge base. Once computed, these individual facet scores are then averaged into a semantic similarity score representing the semantic match between the query and a resource. The facets represented within the dotted border are matched using semantic similarity techniques exploiting the structure of the ontology.

To determine which semantic similarity technique to employ in the semantic matching an evaluation of four different candidate techniques was conducted. This evaluation is described in the next section.

4 EVALUATION

The evaluation is performed as a comparative evaluation by running the algorithm described in Section 3

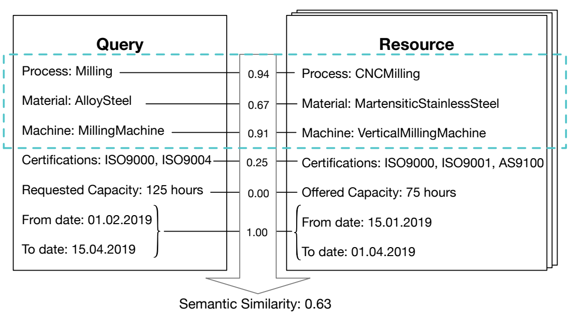


Figure 2: Example of a semantic matching process.

in four different configurations (one for each semantic similarity technique) on a dataset consisting of supplier resources. The dataset was created as follows:

1. We selected a subset of concepts from the Material branch in the ontology (Ferrous Metals).
2. From this subset relevant combinations of material, process and machine were established on the basis of rules specified by domain experts. A rule could for example define that for the material *CarbonSteel*, the process *LaserBeamCutting* and the machine *LaserCuttingMachine* is a valid combination.
3. For the remaining facets (Certificates, Capacity and Calendar Availability) as well as for the supplier data (ID, name, location) we simulated the content using randomly generated input.

Using the above approach we generated 900 resource record instances that were imported into the ontology and used as our dataset. Figure 3 shows an excerpt of two resource records from the test data.

Each of the four algorithm configurations used one of the following semantic similarity techniques:

- Wu-Palmer (Wu and Palmer, 1994). The Wu-Palmer algorithm is an edge-based method that calculates a similarity score by considering the depth of the two concepts to be matched (c_s and c_t), along with the depth of their least common subsumer (lcs):

$$Sim_{wp} = \frac{2 * depth(lcs)}{(depth(c_s) + depth(c_t))} \quad (3)$$

- Resnik (Resnik, 1995). Resnik is an information content-based method that defines the similarity

Supplier	City	Country	Capacity	Certification	Material	Process	Machine	Available from	Available to
Gaylord-Bechtelar	Rakszawa	Poland	125	ISO9001, LEED	CarbonSteel	Shaping	EDMMachine	15.01.19	01.04.19
Walsh LLC	Berlin	Germany	100	AS9000, MIL	AluminumAlloy	CNCMilling	MillingMachine	01.02.19	15.03.19

Figure 3: Format of test data.

between two concepts as the information content of their least common subsumer:

$$Sim_{res} = IC(lcs) \tag{4}$$

- Lin (Lin et al., 1998). Lin extends Resnik by including a calculation of the information content of the two concepts to be matched in addition to the information content of their least common subsumer:

$$Sim_{lin} = \frac{2 * IC(lcs)}{IC(c_s) + IC(c_t)} \tag{5}$$

- Jiang-Conrath (Jiang and Conrath, 1997) propose a hybrid approach that is derived from the edge-based notion by adding the information content as a decision factor. The normalised Jiang-Conrath similarity (Seco et al., 2004) is computed as:

$$Sim_{jc} = 1 - \frac{IC(c_s) + IC(c_t) - 2 * IC(lcs)}{2} \tag{6}$$

Apart from using different similarity techniques, the four configurations used the same approach, allowing to isolate the performance measurement to the similarity technique applied. The evaluation was performed on a machine with Intel Core i7 processor and 16 GB of RAM memory. We generated a composite consumer query that included two sub-queries representing different and randomised variations of the facets. Sub-queries reflect the fact that a consumer may want to request multiple processes in one single query, for example both cutting and assembling metal parts.

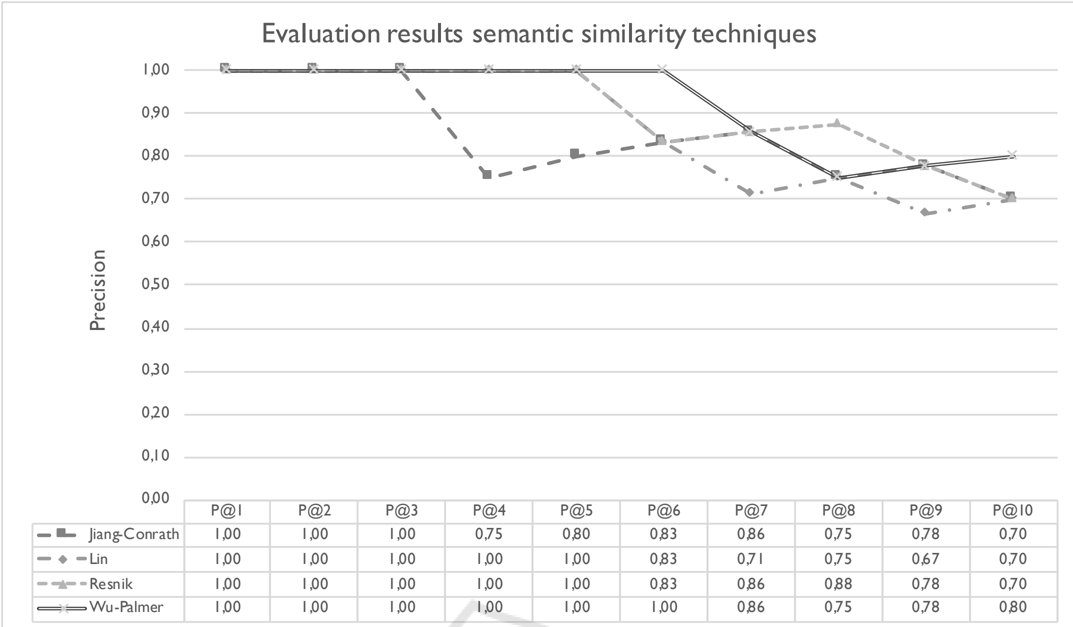
For each of the four configurations, the 10 top resulting hits returned by the algorithm were evaluated for correctness by three domain experts. A majority vote was used to consolidate the evaluation results, hence if two out of three evaluators judged a result as correct, it was finally considered correct. The evaluation measure used was *precision@k* (Elbedweihy et al., 2015), whereby the precision is measured relative to the rank *k* of the search result. For example, precision@3 is 0.67 if 2 out of the three first search results in the ranked list of search results are correct. Since the experts only evaluated the top 10 search results there is no full ground truth alignment from which recall can be measured.

To support the domain experts in their evaluation they were given some extra context information in the form of a hierarchical listing of sub- and superclasses for each of the ontology concepts relevant for each query result. Since the domain experts had little experience with ontologies, such context information is important for the validity and reliability of the evaluation (Cheatham and Hitzler, 2014).

Figure 4 shows the results from the evaluation. As the figure reveals, all four techniques returned only correct search results among the top 3 suppliers. Lin, Resnik and Wu-Palmer also achieves a 100 % precision until the precision@6 threshold where Lin and Resnik return 1 false positive result each while Wu-Palmer maintains its 100 % precision. At precision@10, Wu-Palmer, the edge-based technique, obtains the highest precision of 0.80. These results contradict results from other experiments, e.g., those reported in Resnik (Resnik, 1995) and Seco et al. (Seco et al., 2004), where the information content-based similarity methods perform better than edge-based methods. One possible explanation of this is that many of these experiments base the similarity computation on the WordNet ontology, which describes general knowledge. When domain-specific ontologies are used, as in our case, the results may differ (Pirró, 2009).

There are some validity threats related to the evaluation that should be mentioned. First of all, and in general, the task of assessing the relevance of search results in information retrieval evaluation can be highly subjective (Manning et al., 2010). In this case the queries consisted of multiple parameters and the threshold for determining their collective correctness may vary. For example, should the domain experts weigh some parameters higher than others, or should a search result be deemed correct if 5 out of 7 parameters are considered similar? Furthermore, the experience and knowledge of the domain experts with regard to particular details may also vary. For example, one expert may be aware that a particular machine is applicable to several different materials, while the other experts may not.

Second, the domain experts used different strategies for determining correct versus false search results. The first domain expert required that both sub-queries were fulfilled by a supplier's offered resources in order to state a correct result. The second domain expert considered a search result as correct as long as the resources offered by a supplier satisfied one out of the two sub-queries. As did the third domain ex-



	P@1	P@2	P@3	P@4	P@5	P@6	P@7	P@8	P@9	P@10
Jiang-Conrath	1,00	1,00	1,00	0,75	0,80	0,83	0,86	0,75	0,78	0,70
Lin	1,00	1,00	1,00	1,00	1,00	0,83	0,71	0,75	0,67	0,70
Resnik	1,00	1,00	1,00	1,00	1,00	0,83	0,86	0,88	0,78	0,70
Wu-Palmer	1,00	1,00	1,00	1,00	1,00	1,00	0,86	0,75	0,78	0,80

Figure 4: Evaluation Scores.

pert, however he used a strategy whereby the score was graded into 0 (not correct), 1 (correct, but only for one of the sub-queries) and 2 (correct for both sub-queries). The four returned ranked lists of search results contained mostly the same list of suppliers, although ranked differently. In total there where 13 suppliers returned in the four lists of search results returned (out of a total of 575 suppliers included in the dataset). Looking at the consensus of the three domains experts, all three were in perfect agreement on 5 of 13 the suppliers, 2 of the 3 experts agreed that 4 of the 13 suppliers offered relevant resources, while the remaining 4 suppliers were considered relevant by only 1 of the domain experts.

5 CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

This paper has described the development of a semantic matching algorithm that will support matchmaking between supply and demand of manufacturing resources. The algorithm computes a similarity score based on similarity along six facets which are represented both in a customer query and manufacturing resources offered by suppliers. These facets are process, material, machine, certifications, capacity and calendar availability.

The algorithm employs formal descriptions of manufacturing resources provided by the MANUSQUARE ontology for the first three facets. As a step in selecting a technique that can exploit these descriptions, we conducted a comparative evaluation of the four common semantic similarity techniques Wu-Palmer, Resnik, Lin, and Jiang-Conrath. The evaluation was performed by three domain experts, who assessed four ranked lists of search results returned from the algorithm using each of the four techniques.

The results from the evaluation showed that Wu-Palmer, an edge-based technique, obtained the highest precision overall.

5.2 Future Work

As future work, the algorithm will be extended to also consider other types of data from various endpoints. These data will enable matchmaking by combining the semantic matching presented in this paper with direct and inferred trend analyses (e.g., from historical transactions), reputation benchmarking based on user feedback, and analyses of collaboration patterns, to name a few.

Although this paper focused on the terminological part (the TBox) of the ontology, we also want to exploit ABox capabilities once more instance data is added. One idea along these lines is to use instance matching as a means to correct erroneous data.

While the algorithm currently uses an unweighted approach to compute similarity scores, we will investigate relevant weight configurations for the different facets. This should be done in collaboration with representatives from the supply chain industry to ensure that the weights respond to the search strategies used by those with a demand for supply chain resources.

By-products resulting from production could be used in other production processes, but are often regarded as waste. The categorization of by-products across different industrial sectors can lead to new matches that were not thought of before. As part of the future work, we will consider a categorization of by-products for use in semantic matching.

Finally, a forthcoming and more comprehensive evaluation will be conducted. Such an evaluation will include a larger panel of domain experts to assess the results and more concrete evaluation guidelines to reduce the possibility of validity threats promoted by a clearer distinction between correct and false search results.

ACKNOWLEDGEMENTS

The work presented here was part of the project "MANU-SQUARE - MANUfacturing ecoSystem of QUAlified Resources Exchange" and received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 761145.

REFERENCES

Ameri, F. and Patil, L. (2012). Digital manufacturing market: a semantic web-based framework for agile supply chain deployment. *Journal of Intelligent Manufacturing*, 23(5):1817–1832.

Cheatham, M. and Hitzler, P. (2014). Conference v2. 0: An uncertain version of the oaei conference benchmark. In *International Semantic Web Conference*, pages 33–48. Springer.

Elbedweihy, K. M., Wrigley, S. N., Clough, P., and Ciravegna, F. (2015). An overview of semantic search evaluation initiatives. *Journal of Web Semantics*, 30:82–105.

Jaccard, P. (1901). Distribution de la flore alpine dans le bassin des dranses et dans quelques régions voisines. *Bull Soc Vaudoise Sci Nat*, 37:241–272.

Järvenpää, E., Siltala, N., Hylli, O., and Lanz, M. (2018). Product model ontology and its use in capability-based matchmaking. *Procedia CIRP*, 72(1):1094–1099.

Jiang, J. J. and Conrath, D. W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. In *Proceedings of the 10th Research on Computational Linguistics International Conference*, pages 19–33.

Lin, D. et al. (1998). An information-theoretic definition of similarity. In *Icml*, volume 98, pages 296–304. Citeseer.

Manning, C., Raghavan, P., and Schütze, H. (2010). Introduction to information retrieval. *Natural Language Engineering*, 16(1):100–103.

Pirró, G. (2009). A semantic similarity metric combining features and intrinsic information content. *Data & Knowledge Engineering*, 68(11):1289–1308.

Pirró, G. and Talia, D. (2010). Ufome: An ontology mapping system with strategy prediction capabilities. *Data & Knowledge Engineering*, 69(5):444–471.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. *arXiv preprint cmp-lg/9511007*.

Schönböck, J., Altmann, J., Kapsammer, E., Kimmerstorfer, E., Pröll, B., Retschitzegger, W., and Schwinger, W. (2018). A semantic matchmaking framework for volunteering marketplaces. In *World Conference on Information Systems and Technologies*, pages 701–711. Springer.

Seco, N., Veale, T., and Hayes, J. (2004). An intrinsic information content metric for semantic similarity in wordnet. In *Ecai*, volume 16, page 1089.

Sherer, S. (2005). From supply-chain management to value network advocacy: Implications for e-supply chains. *Supply Chain Management: An International Journal*, 10:77–83.

Tversky, A. (1977). Features of similarity. *Psychological review*, 84(4):327.

Wu, Z. and Palmer, M. (1994). Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pages 133–138. Association for Computational Linguistics.