Worldwide Scholarships Spreading

Jean-Rémi Bourguet

Núcleo de Estudos em Modelagem Conceitual e Ontologias (NEMO)
Federal University of Esprito Santo (UFES) – Vitória, ES – Brazil
Dipartimento di Scienze Politiche ed Ingegneria dell'Informazione (POLCOMING)
Università degli Studi di Sassari (UNISS) – Sassari, RAS – Italy

Abstract

With the inexorable expansion of the semantic layer on the Web and its ecosystem of connected applications, the global citizens expect more and more data expositions coming from public activities. The recent developments in knowledge representation and reasoning push public structures to deploy their data warehouses in parallel of classical websites exhibitions. This article presents an infrastructure to spread the descriptions of scholarships. After introducing the major contributions concerning the semantical annotation of materials occurring in recruitment processes, we describe our case study about the strategy of the University of Sassari concerning the expositions of academical grants. Supported by a core and aligned ontology of the domain we present our prototypical architecture to support and gather the spread of scholarships.

Introduction

The *World Wide Web* is in constant metamorphose, as shown during the last years by the emergence of an ecosystem of human-centered applications (see (Gandon 2003)). Since "the dawn of its time", the *Web* has been orchestrated by a consortium called *W3C* in charge to ensure an efficient dissemination of recommendations stabilizing common languages. This foundation is particularly crucial to coordinate service providers such as *Web* browsers and search engines. Concomitantly with these recommendations, a strengthening of a *Web of Data* supporting the *Web of Documents* became gradually apparent (see (Heath and Bizer 2011)) to constitute the *Semantic Web* (Berners-Lee, Hendler, and Lassila 2001). Two trends are now remarkable in characterizing a document with data on the *Web*: the top-down approaches, occurring by a human establishment of a metadata set describing the document; and the bottom-up approaches, classifying and annotating the document through automated processes (e.g. by natural language processing and semantic similarities (Leacock and Chodorow 1998), by content-based image retrievals (Hare et al. 2006)). In cases of a large number of data, some portals opt for data warehouses, also called triples store, equipped with open access to launch reasoning tasks and output results. The last years, these data warehouses have proliferated on the *Web* such that some researchers decided to map them in a diagram called *LOD* (standing for "Linked Open Data") cloud (see (Max Schmachtenberg and Cyganiak 2014)). They also decided to define a set of eligible criteria for their presences inside the network map. These criteria can be a unique resource identifiers requirement for each data (called URI), the respect of certain semantic format (e.g. `RDFS`, `OWL` or `OWL2`), a minimum amount of triples and links from/towards other warehouses, and a machinery to launch reasoning tasks inside the data warehouse (e.g. endpoint for queries in `SPARQL`). The central warehouse of the *LOD* cloud and kernel of this graph (w.r.t. number of triples and links from/towards other warehouses) is called DBpedia (see (Auer et al. 2007)) and is now seen as the figurehead of the *Semantic Web*. The second prominent warehouse "revolving" around the DBpedia nucleus is called Geonames, which stores a large number of data such as latitude, longitude, altitude and population for each referenced location in the world. Over time, many other warehouses have become some of the major repositories of the *Semantic Web*, including for instance FOAF-Profile (Brickley and Miller 2012), FreeBase (Bollacker et al. 2008) or YAGO (Suchanek, Kasneci, and Weikum 2007).

In conjunction with this development of linked data, the expectations of citizens regarding the open data coming from public activities has become immense (see (Janssen, Charalabidis, and Zuiderwijk 2012)). Nevertheless, for a given reasoning, semantic layers can be necessary to deal with a large amount of open data heterogeneously spread on the *Web* (see (Michalowski et al. 2004)). This issue is particularly remarkable concerning the investigations for academic job offers on the *Web*. From now on, we will use the word scholarship to characterize not an erudition as the seminal sense but an open position provided by a public organization for a mission of research and/or teaching. Thus, if citizens ask for more transparent communications aiming to set up fair competitions concerning the scholarships, they also expect to easily find the different available contests in one global "academic competition arena" (Cantwell 2011). But, due to the usage of domestic database, it will probably remain a pipe dream unless organizations start to promote local integrations of the *Semantic Web* languages. Consequently in this paper, we will propose a prototypical architecture to support and gather the spread of scholarships guided by a core and aligned ontology of the domain.

The remainder of the paper is as follows: in section 2 we will address how the literature tackled this issue; in section 3 we will introduce our case study concerning the University of Sassari (UNISS); in section 4 we will introduce our prototype to aggregate the worldwide scholarships spreading.

Related work

The semantic annotation of the recruitments materials is an old issue. For a long time, researchers in knowledge representation and reasoning have been convinced about the potential fallouts of such practices. For instance, according to (Sicilia 2006), more and more organizations should annotate their vacancies (posted on their official websites), using a set of terms from a controlled vocabulary and allowing the job proposals to reach more possible candidates by generating a more transparent labor market. We will describe here in turn the attempts to use the *Semantic Web* technologies for the job offer dissemination.

According to (Bizer et al. 2005), in order to increase the transparency of the labor market, several national agencies for work in Europe such as the *German Federal Employment Office* and the *Swedish National Labour Market Administration* launched their projects for open data integrations concerning the job offers. In these projects, the recommendation was to use a common vocabulary to categorize the employers offerings and to send them towards a central database. `HR-XML` (XML for Human Resources (Allen and Pilot 2001)) was the standard used to spread the jobs offers in a controlled vocabulary. In (Mochol, Oldakowski, and Heese 2004), a prototype is described (without implementation) and an ontology for human resources is developed by translating in `OWL` some job profiles and skills classifications such as the *Standard Occupational Classification System* and the *North American industry Classification System*. The authors introduced the idea that some job portals could perform semantic similarities between job offers and candidates profiles. In (Mochol, Wache, and Nixon 2007), the authors claimed for semantic techniques such that query approximations to improve the accuracy of job search. In (Bai and Robertson 2010), the authors created some `RDFa` semantic annotations inside the `HTML` code of the webpages describing the jobs of the *British Civil Service* in order to conduct experiments on query performance times (e.g. queries finding jobs with a reasonable salary and close to a given location).

The reverse approaches proposing to create semantic annotation models for CVs to facilitate recruitment were also common in the literature. Some examples are detailed hereafter. In (Bourse et al. 2002), the authors proposed to consider a job or a diploma as a subset of skills, using semantic annotations in `RDF`. In (Yahiaoui, Boufaïda, and Prié 2006), the authors extended this approach by proposing an architecture to establish a similarity score between an annotated CV and some skills expectations.

In this article, we will propose an infrastructure to specifically deal with the academic scholarship representations. We forged our approach in the philosophy of the *Linked Universities*[1] project. *Linked Universities* is an alliance of European universities promoting the development of ontologies to support representations of courses and qualifications (see (Stubbs and Wilson 2006; Kauppinen, Trame, and Westermann 2012)), research activities (see (Baglatzi, Kauppinen, and Keler 2011; Sicilia 2013)), academic publications (see (D'Arcus and Giasson 2009)) and scholarships (see (Börner et al. 2012)).

Application Case

The semantic extractions are now recognized as one of the key leverages in knowledge representation. We realized an extraction of the data of DBpedia and Geonames launching a `SPARQL` query on the endpoint of DBpedia. We extracted all the referenced universities and their geographical coordinates (latitudes and longitudes in `WGS 84` format).

SPARQL query in Pyhton on the endpoint of DBpedia.

```
import rdflib
from rdflib import ConjunctiveGraph
from SPARQLWrapper import SPARQLWrapper, N3
q = SPARQLWrapper("http://dbpedia.org/sparql")
q.setQuery("""
   PREFIX dbp: <http://dbpedia.org/ontology/>
   SELECT ?x ?lat ?long
   WHERE
      {?x a dbp:University;
      geo:lat ?lat ;
      geo:long ?long .}
""")
q.setReturnFormat(N3)
results = q.query().convert()
```

The Figure 1 shows the result we obtained from this query through a Mercator projection using the software GEPHI[2](Bastian, Heymann, and Jacomy 2009).

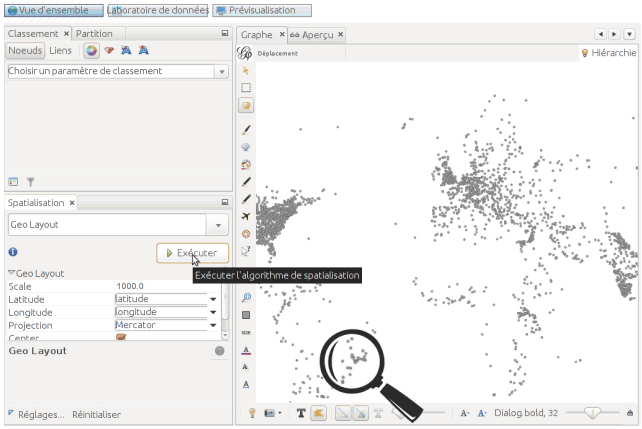


Figure 1: Mercator projection of extracted data

The open data are generally digital data provided by public sources. These data are the most often produced by

[1]http://www.linkeduniversities.org/

[2]https://gephi.org/

local authorities and disseminated in a structured manner under an open license guaranteeing their free access and reuse for citizens. The official portal of UNISS[3] contains in fact an RSS feed (standing for "Rich Site Summary") where the application comes to gather its news. We represent thereafter a fragment of the open data exposure provided by the University of Sassari concerning one of its scholarly activities: the scholarship applications.

Scholarship of UNISS exposed on the official website.

```
<item>
   <title>
      Bando per n. 1 borsa di studio post-
      lauream
   </title>
   <description>
      E' indetto un concorso, per titoli e
      colloquio, per il conferimento di
      una borsa di studio post-lauream per
      motivi di ricerca della durata di 5
      mesi e dell'importo lordo di €10.000
      per lo svolgimento della seguente
      attivita di ricerca: "Calibrazione
      del modello di simulazione EPIC ...
   </description>
   <category>
      Agraria
   </category>
   <link>
      http://hostweb3.ammin.uniss.it/...
   </link>
</item>
```

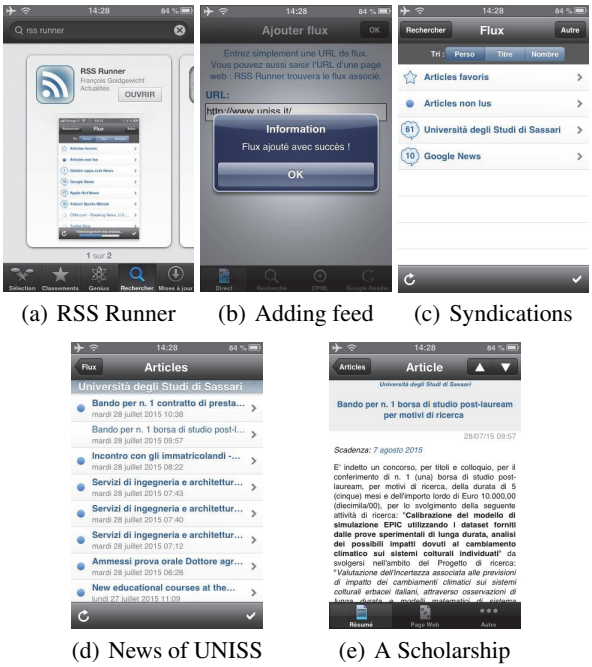
Note that the URL in the markup `<link/>` points towards an XML file more exhaustive that we present below.

Details of the scholarship available with the URI.

```
<pagina tipo="bacheca" id="16704">
   <dataRif_bacheca>
      28/07/2015 14:57:00
   </dataRif_bacheca>
   <dataFine_bacheca>
      07/08/2015
   </dataFine_bacheca>
   <dettaglio numero="1">
      <id_allegati>26931</id_allegati>
      <url_allegati>
         http://www.uniss.it/...
      </url_allegati>
      </dettaglio>
      <dettaglio numero="3">
         <id_allegati>27240</id_allegati>
      <url_allegati>
         http://www.uniss.it/...
      </url_allegati>
   </dettaglio>
</pagina>
```

[3]http://www.uniss.it

If the markups `<title/>` and `<description/>` include important information in natural language format others seem strongly connoted as semantic annotations. The tag `<category/>` serves as a semantic annotation regarding the general scope in which is situated the scholarship. The URL in the markup `<link/>` represents the identity of the scholarship on the *Web* and can therefore also be used as URI. The markup `<dataRif_bacheca/>` concerns the date of issue for the scholarship, `<dataFine_bacheca/>` concerns the deadline for the application and `<url_allegati/>` concerns the annexes which can point towards official decrees (declaration of competition and announcement of the winner). These annexes are important to insure an ethical transparency concerning the scholarship. The Figure 2 is a screenshot of a syndication done in 2015 by the application RSS Runner (available on *Android* and *Apple Store*) with the official URL of UNISS.



(a) RSS Runner (b) Adding feed (c) Syndications

(d) News of UNISS (e) A Scholarship

Figure 2: Portal of UNISS through RSS Runner

In summary, while the scholarships' exposition of the University of Sassari already provides some open data semantically semi-formatted; others (which are sometimes present in the text descriptions) should be reformatted (e.g. salary, referees, expected qualifications). In the next section, we propose an infrastructure to support the expositions and the spread of worldwide scholarships.

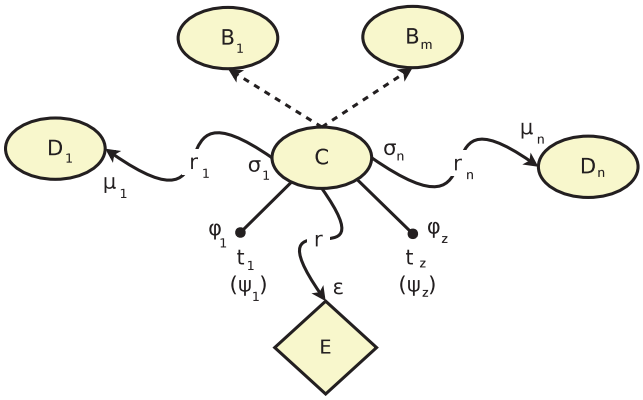
Prototype

One of the major challenge for the *Semantic Web* is to give some leverages to manage a large number of data (big data) across the network having possibly temporal and contextual validities. The alignment of terminological axioms sets (TBoxes) has become an active research track (see for

example (Euzenat and Shvaiko 2013; Ehrig 2006)). As we detailed in the introduction, it also constitutes a required criterion for a presence within the *LOD* cloud. For these reasons, we propose in this section a strategy to align the core ontology we will introduce. This approach will leave wider possibilities for the stakeholders to choose among some terminologies to represent their domestic scholarships.

We define here a graphical interpretation G in $\mathcal{SROIQ}$(Horrocks, Kutz, and Sattler 2006), the fragment of Description Logic underpinning OWL2. We denote $\mathbf{C}$ a set of concepts or classes, $\mathbf{E}$ a set of enumerations, $\mathbf{R}$ a set of object properties or relations, $\mathbf{R_T}$ a set of data properties or attributes, Φ_T a set of datatypes or attributes types, S a set of symbols for Description Logics (see (Baader and Nutt 2003) for their interpretations in first order logic), Ω a set of graphical cardinalities and a function $\ell(\Omega \rightarrow S)$ such that $\ell(*) \mapsto \forall$, $\ell(n) \mapsto =^{\forall} n$, $\ell(*..n) \mapsto \leq^{\forall} n$ and $\ell(n..*) \mapsto \geq^{\forall} n$ (with $n > 0$).

Definition. *Let* $\{C, D_1, \ldots, D_n, B, \ldots, B_m\} \subseteq \mathbf{C}$, $E \in \mathbf{E}$, $\{r, r_1, \ldots, r_n\} \subseteq \mathbf{R}$, $\{t_1, \ldots, t_z\} \subseteq \mathbf{R_T}$, $\{\phi_1, \ldots, \phi_z\} \subseteq \Phi_T$ *and* $\{\mu_1, \ldots, \mu_n, \sigma_1, \ldots, \sigma_n, \psi_1, \ldots, \psi_z\} \subseteq \Omega$, *a graphical interpretation* G *is defined as follows:*



$$(C \sqsubseteq \bigsqcup_i B_i \sqcap \ell(\varepsilon) r.E \sqcap \bigsqcap_j \ell(\mu_j) r_j.D_j \sqcap \bigsqcap_k \ell(\psi_k) t_k.\phi_k)^{\mathrm{G}}$$
$$(D_1 \equiv \ell(\sigma_1) r_1^{-1}.C)^{\mathrm{G}}$$
$$\ldots$$
$$(D_n \equiv \ell(\sigma_n) r_n^{-1}.C)^{\mathrm{G}}$$

We denote $=^{\forall} nr.C$, $\geq^{\forall} nr.C$ and $\leq^{\forall} nr.C$ as the conjunctions of qualified cardinality restrictions and universal restrictions such that $=^{\forall} nr.C \equiv \forall r.C \sqcap = nr.C$, $\geq^{\forall} nr.C \equiv \forall r.C \sqcap \geq nr.C$ and $\leq^{\forall} nr.C \equiv \forall r.C \sqcap \leq nr.C$. The interpretation in OWL2 used above can be summarized as follows: ellipses represent concepts, diamonds represent enumerations, dotted arcs represent inclusions between *concepts*, rounded edges represents *data properties*, curved oriented edges represent *object properties* (labeled by their names). The cardinalities represent the use of *universal* (*), *universal* and *existential* (1..*) and, therefore, *universal* and *cardinality restrictions* (n..*, *..m, n..m). The cardinalities of the *target concepts* are involved with their *object properties* in the *restrictions* of the *source concepts* while the cardinalities of the *source concepts* are involved with their *inverse object properties* in the restrictions of the *target concepts*. We believe that exhaustiveness and contingency in the choices of representations are the crucial expectations for a modeler. This aspect fits perfectly with one of the philosophy promoted by OWL: the open world assumption (Minker 1982). Thus, the joint usage of *universal restrictions* and *cardinality restrictions* allows one to detect any inconsistency while maintaining a freedom to not expose sensitive or unavailable data (e.g. salaries, annexes or deadlines). Some inconsistencies may be handled in cases of violations of cardinality constraints (e.g. the presence of two deadlines) and unexpected *data types* (e.g. deadline unparsable). The Figure 3 presents a graphical representation of our core ontology and its alignments in charge to support and gather the worldwide scholarships spreading.

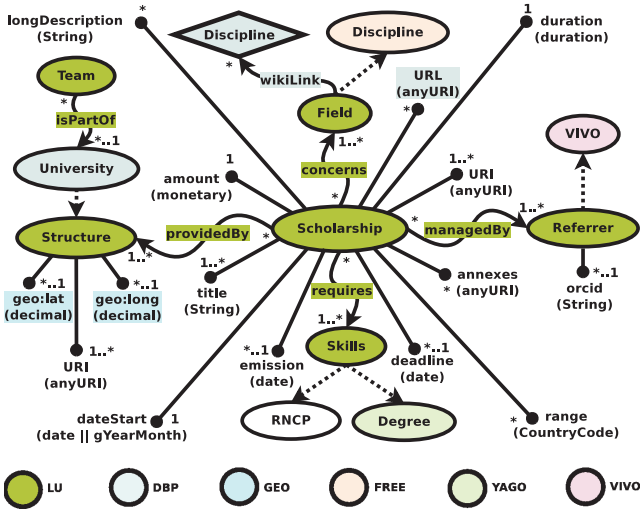


Figure 3: Core ontology and alignments

Once the backbone of our ontology designed, our main intention was to align our concepts with those of the vocabularies already existing.

- LinkedUn (LU) is our core ontology including fundamental concepts of *Scholarship*, *Field*, *Referrer*, *Skill*, *Structure* and *Team*.
- DBpedia (DBP) is the kernel of the *LOD* cloud and constitutes the basis of our syndication mechanism by extracting a set of referenced universities.
- Geonames (GEO) through its alignment with DBpedia supports our extraction of geographical data for the referenced universities.
- Freebase (FREE) is a large and collaborative data warehouse fed by members of the community.
- YAGO (Yet Another Great Ontology) is a warehouse fed by data extracted from WordNet and aligned with DBpedia.
- VIVO is the ontology supporting an open source application (Börner et al. 2012) to represent the academic communities (e.g. researchers, research works, etc.)

While the alignments presented are all founded on class subsumptions, we have also designed alignments supported

by universal restrictions. For example, we introduced in our Tbox the following axiom: if an individual of the class *Field* is related to another individual through the relation wikilink then this individual has to be precisely http://fr.dbpedia.org/page/Liste_des_disciplines_scientifiques, if not an inconsistency may occur. Note that the expressiveness of OWL2 here allows one to use enumerations; leaving the possibility for other lists to figure.

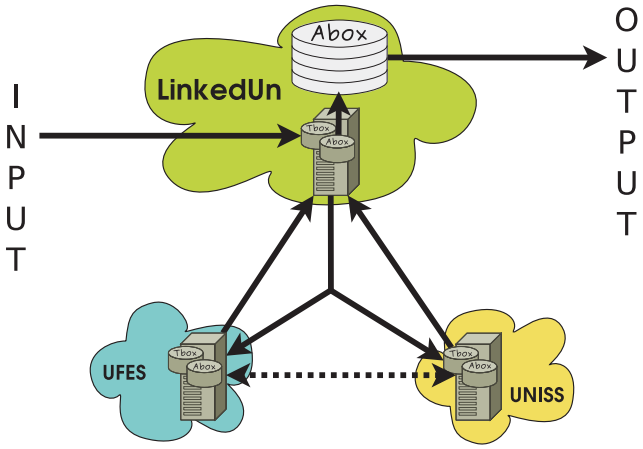


Figure 4: System description

The Figure 4 describes the system description of our prototype, the ontology previously depicted represents the conceptual layer (TBox) of our fluent data warehouse (Abox) fed by different crawlings (straight arrows) through different portals of different universities (e.g. UNISS ABox, UFES Abox, etc.) using possibly some common terminological axioms (dotted arrows).

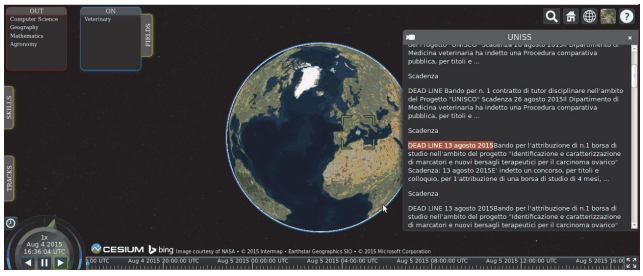


Figure 5: Interface of the prototype LinkedUn

In the end, our prototype includes an interface to perform some queries among the available scholarships (by fields, skills or durations) and to screen the results supported by the API Cesium[4], implemented and depicted in Figure 5.

Conclusion

Over the past decade, the *Web* has become a big digital machinery hovering empirical data with which some stakeholders have already taken advantage. In 2014, the project *Linked Universities* noticed "there are only a few universities currently exposing their public data as linked data, using technologies such as RDF and SPARQL to give direct access to information such as their publications, courses, educational material, etc. [...] and the potential for linked data in education and research goes well beyond the individual benefit for each institution, as this potential can only be achieved through providing cross-university data that can be aggregated, integrated and compared." Following the philosophy promoting by this project, we described in this paper a prototype to support and gather the spread of worldwide scholarship. After an overview of the different propositions used to deal with semantical annotations of materials occurring in a recruitment process, we described the originality of our approach focused on the academic positions. We introduced our case study of the University of Sassari concerning the exposition of scholarship descriptions. We realized some extractions from data warehouses and presented the core ontology supporting our prototype. We designed an ontology with the willingness to leave a maximum of freedom in the scholarship representation.

Finally, we believe in the resurgent profession of ontologist concerning societal, entrepreneurial or governmental activities. The constitution of some ontology-based projects could also push many public structures towards a better integration of the *Semantic Web* technologies following this simple heuristic: if the data are open then the data can be linked. It is undeniable that each country has its own practice and specificity concerning the terminology used to represent their scholarly activities. This fact does not represent an immense obstacle for our approach due to the expressiveness of OWL2 and, more generally, the freedom left in the expression of some domestic vocabularies. Nevertheless, we are strongly convinced that an entity of reference (the European Union, if it concerns a European country or UNESCO for a more global context) should provide an alignment between graduations of the different countries to facilitate the skills elicitation.

Acknowledgments

This work has been made possible by "la Regione Autonoma della Sardegna e Autorità Portuale di Cagliari con L.R. 7/2007, Tender 16 2011, CRP-49656 con il projeto: Metodi innovativi per il supporto alle decisioni riguardanti lottimizzazione delle attività in un terminal container" and by "o EDITAL FAPES/CAPES N° 009/2014 (Bolsa de fixacão de doutores) com a proposta: Melhor integração de tecnologias de representação de conhecimento e raciocínio nas utilizações local e Web".

References

Allen, C., and Pilot, L. 2001. Hr-xml: Enabling pervasive hre-business. Technical report, HR-XML Consortium.

Auer, S.; Bizer, C.; Kobilarov, G.; Lehmann, J.; Cyganiak, R.; and Ives, Z. G. 2007. Dbpedia: A nucleus for a web of open data. In *6th International Semantic Web Conference*

[4]https://cesiumjs.org/

and 2nd Asian Semantic Web Conference, volume 4825 of *LNCS*, 722–735. Springer.

Baader, F., and Nutt, W. 2003. Basic description logics. In *Description Logic Handbook*, 43–95. Cambridge University Press.

Baglatzi, A.; Kauppinen, T.; and Keler, C. 2011. Linked science core vocabulary specification.

Bai, X., and Robertson, D. 2010. Service choreography meets the web of data via micro-data. In *AAAI Spring Symposium: Linked Data Meets Artificial Intelligence*. AAAI.

Bastian, M.; Heymann, S.; and Jacomy, M. 2009. Gephi: An open source software for exploring and manipulating networks. In *ICWSM*. The AAAI Press.

Berners-Lee, T.; Hendler, J.; and Lassila, O. 2001. The semantic web. *Scientific American* 284(5):34–43.

Bizer, C.; Heese, R.; Mochol, M.; Oldakowski, R.; Tolksdorf, R.; and Eckstein, R. 2005. The impact of semantic web technologies on job recruitment processes. In *Internationale Tagung Wirtschaftsinformatik 2005*, 1367–1382. Physica-Verlag.

Bollacker, K.; Evans, C.; Paritosh, P.; Sturge, T.; and Taylor, J. 2008. Freebase: A collaboratively created graph database for structuring human knowledge. In *International Conference on Management of Data*, 1247–1250. New York, NY, USA: ACM.

Börner, K.; Conlon, M.; Corson-Rikert, J.; and Ding, Y. 2012. *VIVO: A Semantic Approach to Scholarly Networking and Discovery*. Synthesis Lectures on the Semantic Web. Morgan and Claypool.

Bourse, M.; Harzallah, M.; Leclère, M.; and Trichet, F. 2002. COMMONCV: modeling the competencies underlying a curriculum vita. Technical report, Institut de Recherche en Informatique de Nantes.

Brickley, D., and Miller, L. 2012. Foaf vocabulary specification 0.98. *Namespace document* 9.

Cantwell, B. 2011. Transnational mobility and international academic employment: Gatekeeping in an academic competition arena. *Minerva* 49(4):425–445.

D'Arcus, B., and Giasson, F. 2009. Bibliographic ontology specification.

Ehrig, M. 2006. *Ontology alignment: bridging the semantic gap*, volume 4. Springer Science & Business Media.

Euzenat, J., and Shvaiko, P. 2013. *Ontology Matching, Second Edition*. Springer.

Gandon, F. L. 2003. Combining reactive & deliberative agents for complete ecosystems in infospheres. In *International Conference on Intelligent Agent Technology*, 297–303. IEEE Computer Society.

Hare, J. S.; Sinclair, P. A. S.; Lewis, P. H.; Martinez, K.; Enser, P. G. B.; and Sandom, C. J. 2006. Bridging the semantic gap in multimedia information retrieval: Top-down and bottom-up approaches. volume 187 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Heath, T., and Bizer, C. 2011. *Linked Data: Evolving the Web into a Global Data Space*. Synthesis Lectures on the Semantic Web. Morgan & Claypool Publishers.

Horrocks, I.; Kutz, O.; and Sattler, U. 2006. The even more irresistible SROIQ. In *10th International Conference on Principles of Knowledge Representation and Reasoning*, 57–67. AAAI Press.

Janssen, M.; Charalabidis, Y.; and Zuiderwijk, A. 2012. Benefits, adoption barriers and myths of open data and open government. *IS Management* 29(4):258–268.

Kauppinen, T.; Trame, J.; and Westermann, A. 2012. Teaching core vocabulary specification.

Leacock, C., and Chodorow, M. 1998. Combining local context and wordnet similarity for word sense identification. *WordNet: An Electronic Lexical Database* 265–283.

Max Schmachtenberg, Christian Bizer, A. J., and Cyganiak, R. 2014. Linking open data cloud diagram 2014.

Michalowski, M.; Ambite, J. L.; Knoblock, C. A.; Minton, S.; Thakkar, S.; and Tuchinda, R. 2004. Retrieving and semantically integrating heterogeneous data from the web. *IEEE Intelligent Systems* 19(3):72–79.

Minker, J. 1982. On indefinite databases and the closed world assumption. In *6th Conference on Automated Deduction*, volume 138 of *LNCS*, 292–308. Springer.

Mochol, M.; Oldakowski, R.; and Heese, R. 2004. Ontology based recruitment process. In *der 34. Jahrestagung der Gesellschaft für Informatik*, volume 51 of *LNI*, 198–202. GI.

Mochol, M.; Wache, H.; and Nixon, L. J. B. 2007. Improving the accuracy of job search with semantic techniques. In *10th International Conference of Business Information Systems*, volume 4439 of *LNCS*, 301–313. Springer.

Sicilia, M.-A. 2006. *Competencies in Organizational E-Learning: Concepts and Tools*. Information Science Publishing.

Sicilia, M.-A. 2013. The common european research information format ontology cerif ontology 0.2.

Stubbs, M., and Wilson, S. 2006. eXchanging course-related information: a UK service-oriented approach. *TENC: Publications and Preprints*.

Suchanek, F. M.; Kasneci, G.; and Weikum, G. 2007. Yago: A core of semantic knowledge. In *Proceedings of the 16th International Conference on World Wide Web*, 697–706. New York, NY, USA: ACM.

Yahiaoui, L.; Boufaïda, Z.; and Prié, Y. 2006. Semantic annotation of documents applied to e-recruitment. In *3rd Italian Semantic Web Workshop of Semantic Web Applications and Perspectives*, volume 201 of *CEUR Workshop Proceedings*. CEUR-WS.org.