20. Ontologies

Roberto Navigli*
Sapienza University of Rome

Abstract

This chapter is about ontologies, that is, knowledge models of a domain of interest. We introduce ontologies, view them from the perspective of several fields of knowledge, and present existing ontologies and the different tasks of ontology building, learning, matching, mapping and merging. We also review interfaces for building ontologies and the knowledge representation languages used to implement them. Finally, we discuss the different ways of evaluating an ontology and the applications in which it can be used.

Keywords: Ontologies, Knowledge representation, Lexical semantics, Semantic Web

20.1 Introduction

In computational linguistics and computer science, an **ontology** is a formal representation of knowledge. Since ancient times human beings have constantly searched for new ways to express and encode their knowledge. Nevertheless, until recently this knowledge has overwhelmingly been represented by means of informal tools, such as natural language and pictures. Today, however, with the advent of computers – and the Web era – it is becoming increasingly clear that formally encoding knowledge would make possible a new generation of the Web to enable information processing at the meaning level.

* Dipartimento di Informatica, Sapienza Università di Roma, Viale Regina Elena, 295b, 00161 Roma, Italy, E-mail: navigli@di.uniroma1.it.

In fact, ontologies are about meaning. A popular definition for an ontology is "a formal specification of a shared conceptualisation" (Gruber 1993). This definition makes it clear that we need to represent formally and explicitly our model of the knowledge we are interested in (typically, a domain) and that this model should be agreed among users, experts, communities, etc. In other words, we can say that an ontology is a set of definitions in a formal language for concepts that describe the world of interest, including the relationships that connect these concepts.

So, ontologies are about formalising knowledge. But how formal and explicit are ontologies? This question can be answered by comparing the degree of formalisation of ontologies with that of other resources such as terminologies, glossaries, thesauri and taxonomies. As can be seen in Figure 1, the degree of formalisation constantly increases from the least to the most formalised knowledge resource: **unstructured text** – just a string of text with no additional structure; **terminology** – a set of terms expressing concepts for the domain of interest (e.g. hotel, room, tourist, etc.); **glossary** – a terminology with textual definitions for each term (e.g. "an establishment that provides short-term lodging" as definition of hotel); **thesaurus** – which provides information about relationships between words, like synonyms (e.g. motel is a synonym of motor hotel) and antonyms (e.g. ugly is an antonym of beautiful); **taxonomy** – a hierarchical classification of concepts (e.g. a motel *is-a* hotel); **ontology** – a fully-structured knowledge model, including concepts, relations of various kinds and, possibly, rules and axioms.

20.2 Anatomy of an Ontology

20.2.1 Building blocks of an Ontology. An ontology is composed of the following building blocks:

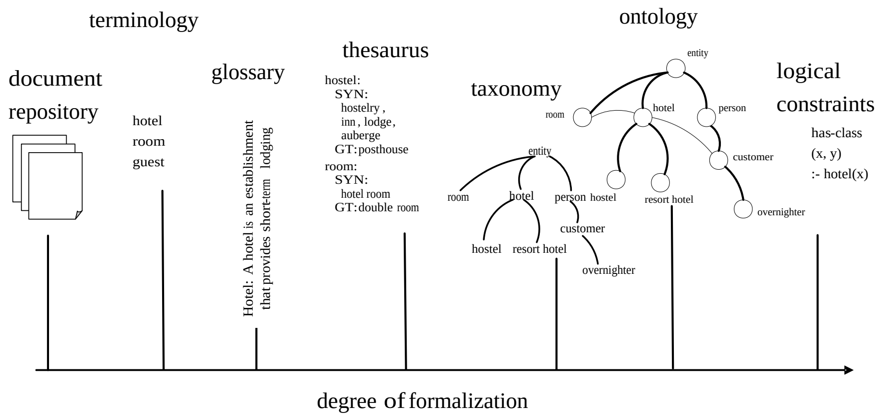

Figure 1
The different degrees of formalisation: from unstructured textual content to ontologies.

- **Concepts** (or **classes**), which represent the basic units of an ontology. A concept identifies a meaning that the ontology creators want to include in their representation of the domain. If an ontology is **lexicalised**, a concept is associated with one or more terms that express it by means of language. For instance, given the *car* concept in the automobile domain, synonyms such as *car*, *automobile*, *motorcar* are typically associated with that concept. In knowledge representation languages (cf. Section 20.8), a **TBox** (terminological box) includes the set of concepts of an ontology.
- **Instances** (or **individuals** or **objects**), which represent the ground level of the ontology. These are objects of the real world, such as an existing car license plate number in the domain of interest (e.g. LO108ST). In knowledge representation languages (cf. Section 20.8), an **ABox** (assertional box) includes the set of instances of an ontology.
- **Relations**, which connect concepts and invididuals to one another. Among the most popular (and relevant) ontological relations we mention:

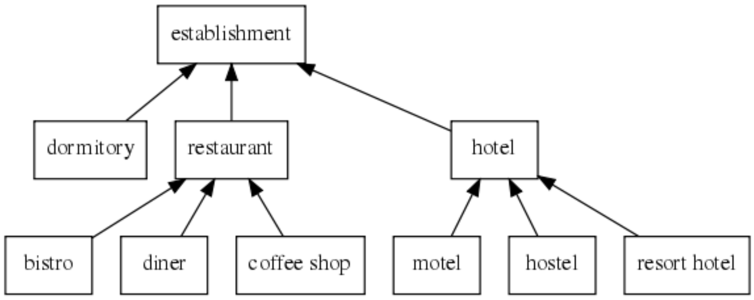

Figure 2
An excerpt of taxonomy.

- The **is-a** (or **type-of** or **subclass-of**) relation (also called **hypernymy** in lexical ontologies). An ontology whose only relations between concepts are of this kind is called a **taxonomy**, an excerpt of which is shown in Figure 2 (e.g. a motel *is-a* hotel according to the taxonomy).
- The **instance-of** relation, which connects each instance to the concept that represents its abstract counterpart. For example, LO108ST *instance-of* car license plate number.
- The **has-a** (or **has-part**) relation (also called **meronymy** in lexical ontologies, e.g. a hotel *has-part* hotel room).

Relation labels do not always have semantics which are clearly established (e.g. in standardised ontology languages, cf. Section 20.8). As mentioned above, the same relation may have different names (e.g. *is-a* vs. *subclass-of*), whereas different relations may have the same name (e.g. *produces* as in Pink Floyd *produces* The Dark Side of the Moon vs. Fiat *produces* Car).

Ontologies can further include the following elements:

- **Attributes** (or **properties**), which represent relations intrinsic to specific concepts (e.g. qualities such as color, measures such as a person's height and name).
- **Restrictions** on relations (e.g. the *has-parent* relation can connect only instances of the human concept).
- **Rules** and **axioms**: assertions in a logical form that encode the overall theory that the domain ontology describes.

20.2.2 Sections of an Ontology. Ontologies are composed of the following sections:

- An **upper ontology** (or **top ontology**), that encodes high-level concepts and relations, which do not belong to a specific domain of interest. Upper ontologies aim to enable semantic interoperability between different ontologies by providing the most general concepts structured in a hierarchy and optionally associating general rules and axioms about those concepts. Existing upper ontologies – introduced in Section 20.4.1 – include SUMO, the WordNet top ontology, and the Cyc upper ontology.
- A **middle** or **general-purpose ontology** that encodes general concepts (units of measurement, spatial and temporal relations, communication, mental and physical objects, etc.) which allow connections to be made between more specific concepts usually encoded in a domain ontology. Existing middle ontologies – introduced in Section 20.4.2 – include WordNet and Cyc.
- A **domain ontology** that instead models concepts, individuals and relations about the knowledge domain of interest. Different domain

ontologies can either use the same upper/middle ontology or provide a mapping to a common upper/middle ontology, thus enabling interoperability between them. Existing domain ontologies – introduced in Section 20.4.3 – include UMLS and the Gene Ontology.

- An **application ontology** – an ontology developed for a specific use or application focus. Its scope is typically defined on the basis of use cases that can be used to test the ontology. Application ontologies depend both on domains and on a specific task of interest, and are typically used when crossing domains (e.g. the geospatial field).

The different sections of an ontology grow in size approximately from 10-100 for an upper ontology to thousands or millions for a domain or application ontology.

An important principle behind ontologies – which also justifies the above modularisation – is *reuse*: applications do not need to reinvent the wheel, as knowledge has most likely already been encoded in one or more existing ontologies. The ideal use of an ontology is to plug it into an application of interest and use that structured knowledge for a specific purpose (reasoning, semantic processing, etc.). In fact, ontologies are the opposite of reinventing the wheel, similarly to what happens in software engineering (cf. Section 20.3.5).

20.3 Ontologies under different lenses

20.3.1 Computer science vs. Philosophy. Humans have long studied abstract ways to model reality. This kind of philosophical study, namely the study of being, is called *ontology*. In fact, ontology studies the nature of being and existence, together with the basic categories of being and their relations. The jump to computer science is short. If we need to formalise and model the knowledge of a specific domain, we need a "formal specification of a shared conceptualisation" (Gruber 1993), i.e. an ontology. In computer science, an ontology must be *formal*, because it must be encoded and processed as a data structure in a computer, and it must model a conceptualisation that is *shared*, because an ontology is aimed at enabling interoperability, thus it needs to encode knowledge in a way that is shared by domain experts and users. Ontologies are used in computer science because they provide a structured data model for knowledge that can be used and processed within computer programs.

20.3.2 Ontologies and the Semantic Web. Ontologies are the backbone of the **Semantic Web** – a vision of the Web in which computers can semantically process and interpret the information provided on the World Wide Web. In fact, the knowledge modelled by one or more ontologies can be used to semantically annotate Web pages, perform semantic search, create agents that understand user needs or participate in a dialogue among remote applications, etc. (see Section 20.10 for an illustration of different applications of ontologies). In this sense, ontologies are the common ground for performing any kind of semantically-orientated task aimed at implementing the Semantic Web and making applications interoperate. To give a clearer idea of the part ontologies play in this vision, in Figure 3 we reproduce the so-called Semantic Web layer cake, which illustrates the architecture of the Semantic Web. On the bottom of the cake we have strings used to identify a name or a resource on the Internet (**Uniform Resource Identifiers** or **URIs**) and character encodings (e.g. **Unicode**). Immediately on top of these, we have the **eXtended Markup Language** (**XML**[1]), used to encode documents in a structured machine-readable format. XML is used to build the instance level (encoded as (subject, relation, object) triples in **RDF**) and the taxonomic level of ontologies (written in **RDFS**). Full ontologies find their place alongside logical rules (expressed in the Rule Interchange Format or RIF (Kifer 2008)) and on top of taxonomies. A specific

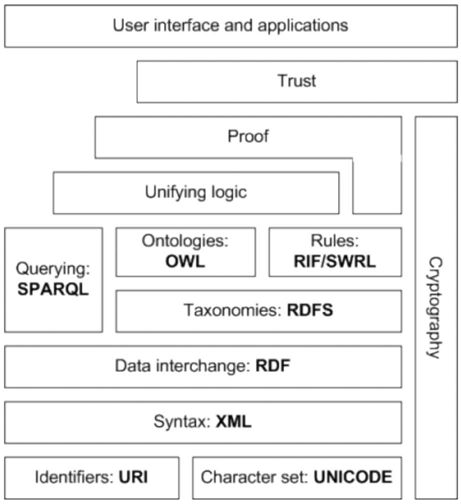

Figure 3
The Semantic Web layer cake.

query language is used for ontologies, namely **SPARQL**[2]. The topmost levels of the cake deal with the logical and semantic validation of ontologies. In the last few years, formal ontologies have given way to a more lightweight, distributed representation of knowledge, called **Linked Data (LD)**, leading to the creation of the so-called Linked (Open) Data Cloud.[3] The languages used to represent ontologies and linked data (RDF, RDFS, OWL, etc.) are discussed in Section 20.8.

20.3.3 Ontologies and the Lexicon. A **computational lexicon** is a structured lexical resource that encodes meanings in terms of the words that express them. Computational lexicons are the general-purpose counterpart of domain ontologies; in this sense they might be considered upper or middle ontologies. However, these lexicons also contain a large amount of domain information (i.e. concepts and relations about various different domains). The prototypical example of a computational lexicon is WordNet (Fellbaum 1998), which contains over 117,000 concepts, named **synsets** (i.e. synonym sets). We can view a computational lexicon as a **lexical ontology**, that is, an ontology whose concepts are associated with the terms used to express them.

Although lexicons and ontologies have much in common, there is an inherent distinction to keep in mind: the former are linguistic objects (i.e. they depend on a natural language), while the latter are non-linguistic and represent the relations between sets of objects or abstractions in the world of interest (see (Hirst 2009) for a discussion).

20.3.4 Ontologies and Graphs. A typical view of ontologies is that of **semantic networks**. Semantic networks are **directed** or **undirected graphs** $G = (V, E)$ whose set of vertices V represents concepts and whose edges E are semantic relations between concepts. For example, WordNet is a semantic network where vertices are synsets and edges are relations such as hypernymy (*is-a*), meronymy (*part-of*), etc. We show an excerpt of the WordNet semantic network in Figure 4. Note that semantic networks must be distinguished from **conceptual graphs**, a logical formalism used to represent statements in first-order predicate logic (see (Sowa 2000)).

An important issue in encoding hierarchical taxonomies is the *single vs. multiple inheritance* question: should a concept be constrained to be a subclass of only one concept (*single inheritance*) or should it be allowed to be a subclass of one or more concepts (*multiple inheritance*)? For instance, is a beverage a kind of food or liquid? Probably both. To cope with this need, ontologies such as WordNet allow some limited form of multiple inheritance. However, inconsistencies such as the **Nixon diamond** problem (Reiter and Criscuolo 1981) can arise. Assume our ontology states that: 1) a Quaker *is-a* pacifist, 2) a Republican *is-a* hawk (i.e. is *not* a pacifist), 3) Nixon *instance-of* Quaker, 4) Nixon *instance-of* Republican. We have a clear contradiction here: is Nixon a pacifist or a hawk? A possible solution to this problem is the use of concept **facets** that implement

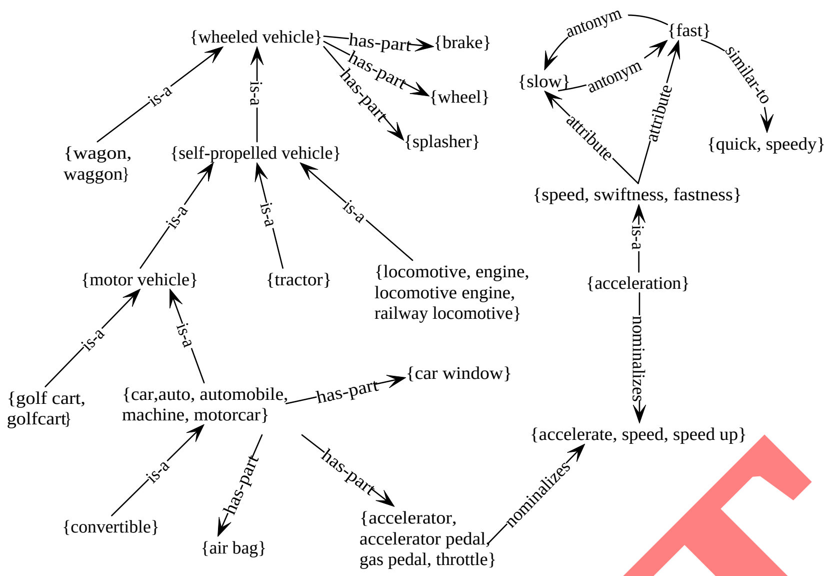

Figure 4
An excerpt of the WordNet semantic network.

restrictions on relations or properties. For instance, birds have the property of being able to fly, but penguins (that are a *type-of* bird) do not.

In terms of the ontology graph, assuming a unique root exists, a single inheritance taxonomy is a **tree**, whereas a multiple inheritance taxonomy is a **semilattice**, that is, a partially ordered set with a least upper bound for any nonempty finite subset of concepts.

The graph structure view of an ontology can be used to perform a variety of operations, such as determining the semantic similarity between pairs of concepts (e.g. (Jiang and Conrath 1997; Leacock and Chodorow 1998; Pilehvar, Jurgens, and Navigli 2013)), performing Word Sense Disambiguation (Navigli and Lapata 2010), **entity linking** (Ferragina and Scaiella 2010; Moro, Raganato, and Navigli 2014), etc.

20.3.5 Ontologies and Software Engineering. As mentioned in Section 20.2.2, one of the main purposes of ontologies is reuse. We encode knowledge in an ontology to share and reuse it. Designing and implementing a good ontology is similar to designing and implementing good software. In fact, ontologies are a special piece of software. Thus, it is natural to compare ontology construction with software engineering. It has been argued that a software engineering process, such as the Unified Process, can be used for building ontologies as well (De Nicola, Missikoff, and Navigli 2009). Furthermore, ontology design patterns[4] have been identified that can be employed as the building blocks of the ontology engineering process (Gangemi and Presutti 2009). Patterns exist at various different levels, such as the content level (e.g. parts of a concept), the lexico-syntactic level (e.g. providing synonyms to express a concept), the logical level (e.g. partitions of concepts), etc. Conversely, ontologies can be used during the software engineering process to describe requirements and formally represent the knowledge these requirements encode, so as to make specifications more precise and easily traced and maintained. Ontologies can also be used to describe the functionalities of software components, thus facilitating component reuse. More uses of ontologies in software engineering are possible, e.g. in supporting the coding and deployment phases (see (Seedorf, Informatik, and Mannheim 2006) for a thorough survey on the topic).

20.4 Existing ontologies

We now review well-known upper (Section 20.4.1), middle (Section 20.4.2) and domain (Section 20.4.3) ontologies. We stress again that these can be used in various different combinations.

20.4.1 Upper Ontologies.

- **Suggested Upper Merged Ontology (SUMO)**[5]: a foundational ontology created for a variety of information processing tasks (Pease, Niles, and Li

2002). SUMO includes more than 1,000 concepts and about 4,000 relations between them. It was created by merging a number of existing upper-level ontologies, including abstract ones (Sowa 2000; Borgo, Guarino, and Masolo 1996) and more concrete ones developed at Stanford KSL and ITBM-CNR. SUMO also includes a mid-level ontology and a variety of domain ontologies, providing several thousand formal axioms.

- **WordNet top ontology** (Fellbaum 1998): the upper part of the WordNet noun taxonomy, including the 51 most general concepts or **unique beginners** (such as entity, physical object, abstraction, group, relation, measure, etc.).
- **EuroWordNet top ontology** (Vossen 1998): an ontology consisting of 63 high-level concepts. Concepts are classified as first-order (concrete entities that can be perceived by the senses), second-order (static and dynamic situations, such as properties or relations) and third-order (unobservable concepts, such as ideas, plans, etc.).
- **Cyc Upper ontology**: an ontology containing about 3,000 general concepts that make up the upper part of the Cyc ontology (Lenat 1995).
- **Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)** (Masolo et al. 2003): a cognitively-biased taxonomy of ontological categories underlying natural language and commonsense. Basic categories include endurants (e.g. physical objects), perdurants (e.g. events), qualities (e.g. spatial locations) and abstracts (facts, sets, etc.).
- **CRM CIDOC** (Crofts et al. 2010): an upper ontology aimed at enabling the integration and exchange of cultural heritage information.

20.4.2 Middle Ontologies. Most of the middle ontologies available online are general-purpose, in that they provide all the semantics needed to later attach further domain-specific concepts:

- **WordNet**[6] (Fellbaum 1998): a semantic network of English organised according to psycholinguistic principles. Although it is a general-purpose ontology, some parts of the WordNet taxonomy concern specific domains. WordNet concepts have also been explicitly marked with domain labels by Magnini and Cavaglià (2000), including a hierarchy of domains (an excerpt of which is shown in Figure 5).
- **BabelNet**[7] (Navigli and Ponzetto 2012): a very large, wide-coverage multilingual semantic network made up of 13.9 million concepts and named entities (as of version 3.0). The network is automatically constructed by means of the seamless integration of lexicographic and encyclopedic knowledge from WordNet, Wikipedia, Wikidata, OmegaWiki, Wiktionary and the Open Multilingual WordNet (Bond and Paik 2012). Concepts are lexicalised in many languages and relations between concepts include those from WordNet (e.g. *is-a* and *part-of*) and unlabeled relatedness relations harvested from Wikipedia.
- **The Wikipedia Bitaxonomy**[8] (Flati et al. 2014, WiBi): a very large automatically-integrated taxonomy of English Wikipage pages and categories. In contrast to other taxonomies, like that of WordNet, WiBi covers encyclopedic knowledge (e.g. Zucchero Fornaciari *is-a* songwriter) and is integrated into BabelNet (starting with version 3.0).

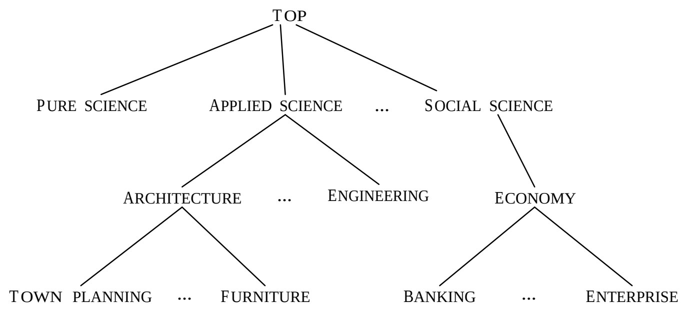

Figure 5
An excerpt of the WordNet domain labels taxonomy.

- **Cyc**[9] (Lenat 1995): a wide-coverage ontology of commonsense knowledge. The current open-source version of Cyc (**OpenCyc**) includes almost 50,000 concepts and more than 300,000 relations between concept pairs.
- **Yet Another Great Ontology (YAGO)**[10] (Suchanek, Kasneci, and Weikum 2007; Hoffart et al. 2013): a large ontology built automatically from Wikipedia. YAGO includes over 10 million named entities (such as persons, cities and organisations) and about 120 million relations between entities (e.g. AlbertEinstein *has-won-prize* NobelPrize).
- **DBPedia**[11] (Auer et al. 2007): a lightweight, cross-domain ontology, manually created from Wikipedia infoboxes. The ontology contains around 4.5 million entities, including places, persons, works, species, organisations and buildings.
- **Omega**[12] (Philpot, Hovy, and Pantel 2005): a terminological ontology containing about 120,000 concepts obtained by reorganising two large ontologies, namely WordNet and Mikrokosmos.

20.4.3 Domain Ontologies.

- **Unified Medical Language System (UMLS)**[13] (McCray and Nelson 1995), which includes a semantic network providing a categorisation of medical concepts.
- **Systematised Nomenclature of Medicine Clinical Terms (SNOMED-CT)**[14], whose ontology includes a core set of over 364,000 health care concepts organised into taxonomic hierarchies.
- **Gene Ontology**[15] (The Gene Ontology Consortium 2008): a collaborative effort in the field of bioinformatics to standardise the representation of gene and gene attributes in a domain ontology. The ontology covers three domains: cellular components, molecular function and biological process.
- **PRotein Ontology (PRO)**[16] (Natale et al. 2006): a formal representation of proteins, including their formalisation as concepts and the relationships between them. The ontology includes a "sub-ontology of proteins based on evolutionary relatedness and a sub-ontology of the multiple protein forms produced from a given gene".
- the **North American Industry Classification System (NAICS)** and the former **Standard Industrial Classification (SIC)**: domain taxonomies aimed at classifying industrial services.

20.5 Ontology Building vs. Ontology Learning

20.5.1 Building. Ontologies can be created manually through the efforts of domain experts, a task referred to as **ontology building** or **ontology construction**. This manual process typically involves the following steps:

- **Requirements and Analysis**: information resources are collected and experts are asked to define the terms that formally describe concepts in the domain of interest.
- **Design**: the conceptual organisation of the ontology is designed. Which are the concrete concepts (possibly instances) and relations to encode?
- **Implementation**: The ontology is written in a specific language, e.g. RDF or OWL (cf. Section 20.8).
- **Test**: inconsistencies of different kinds are reconciled and the general consistency of the ontology is checked.

Finally, the ontology is released. The ontology building process can be iterated to further refine the ontology. Different methodologies have been proposed that establish guidelines for ontology building, including:

- **METHONTOLOGY** (Fernandez-Lopez, Gomez-Perez, and Juristo 1997): a methodology for building ontologies either from scratch or via a re-engineering process. The methodology clearly specifies the steps to perform to build the ontology;
- **On-To-Knowledge** (Sure et al. 2003): a knowledge engineering methodology that consists of 5 phases: feasibility study, kick-off, evaluation, refinement, application and evolution.
- the **Unified Process for ONtology building (UPON)** (De Nicola, Missikoff, and Navigli 2009): an ontology development methodology stemming from the Unified Process for software engineering.

20.5.2 Learning. The manual construction of ontologies is costly and usually requires the agreement of the domain experts involved in the process. This issue can be addressed by means of **ontology learning**, i.e. techniques aimed at (semi-)automatically acquiring an ontology. If the instance level is involved (i.e. real-world individuals), the automatic acquisition process is called **ontology population**. Ontology learning and population has the advantage of reducing not only the costs of construction but also those of maintenance, which often has to be carried out for several years.

The steps required to learn an ontology are linguistically-grounded, in the sense that terms, relations and axioms are extracted from domain texts with Natural Language Processing techniques. The following steps are usually performed:

- **Term extraction**: this task consists of the automatic acquisition of domain terms from raw text (e.g. hotel, motel, motor hotel, etc. in the tourism domain). Techniques range from the use of TF-IDF to more complex measures such as specificity and cohesion (Park, Byrd, and Boguraev 2002), domain consensus and relevance (Navigli and Velardi 2004), etc. (see also Chapter 38). This step might also include the identification of synonyms (e.g. motel and motor hotel) with corpus-based (Rapp 2003), lexicon-based (Jarmasz and Szpakowicz 2003) and hybrid approaches (Turney et al. 2003). The resulting sets of synonyms represent the ontology concepts. Glosses, i.e. textual definitions, can be further harvested and associated with terms (Velardi, Navigli, and D'Amadio 2008; Navigli and Velardi 2010).
- **Taxonomy learning**: concepts are then structured in a taxonomic hierarchy. This step is performed with the aid of lexico-syntactic patterns

(Hearst 1992), also combined with graph-based methods (Kozareva and Hovy 2010b), taxonomy restructuring based on word sense disambiguation (Navigli and Velardi 2004), clustering techniques (Cimiano, Hotho, and Staab 2005), and hypernym extraction from textual definitions (Velardi, Faralli, and Navigli 2013).

- **Relation learning**: next, non-taxonomic relations are learned (e.g. *part-of*, *location*, *purpose*), possibly including domain-specific relations. Typically, semantic relations are harvested from text by means of statistical measures of word co-occurrence (Maedche and Staab 2000; Hasegawa, Sekine, and Grishman 2004; Pantel and Pennacchiotti 2006), the use of regular expressions (Navigli and Velardi 2008), recursive lexico-syntactic patterns (Kozareva and Hovy 2010a), and Open Information Extraction techniques (Fader, Soderland, and Etzioni 2011; Moro and Navigli 2013) to model the surface meaning of semantic relations.
- **Learning of facts and axioms**: finally, facts and axioms can be automatically extracted from text. Approaches include the automatic acquisition of generalised extraction patterns and similarity-based fact ranking (Paşca et al. 2006), the analysis of textual definitions (Völker, Hitzler, and Cimiano 2007), the use of linguistic patterns to extract facts from the Web (Etzioni et al. 2004) and iterative fact learning based on a set of knowledge extraction components (Carlson et al. 2010). In order to prune out noise, the set of extracted facts can be ranked by means of the PageRank algorithm (Jain and Pantel 2010).

Well-known ontology learning systems include: OntoLearn (Navigli and Velardi 2004), OntoLT (Buitelaar, Olejnik, and Sintek 2004), TextToOnto (Maedche and Volz 2001), Text2Onto (Cimiano and Völker 2005) and, more recently, OntoLearn Reloaded (Velardi, Faralli, and Navigli 2013).

20.5.3 Maintenance. Finally, we mention here an issue that is very important regardless of whether an ontology has been created manually or automatically: **ontology maintenance**. Maintaining ontologies is the task concerned with keeping them up-to-date, performing versioning and avoiding incompatibilities with older versions. Similarly to what happens with software, maintaining an ontology is a hard task. However, the task can be partially automatised by means of algorithmic techniques (e.g. by pruning and refining ontologies (Maedche and Volz 2001)).

20.6 Ontology Matching, Mapping and Merging

It is not infrequent that many ontologies exist for the same domain. It might also happen that several ontologies for different domains have to be used within the same application and have a considerable overlap (for instance, ontologies for the domains of business and music – with many concepts in common). Finally, different versions of the same ontology might be produced. In all these cases, it is desirable to find correspondences between entities of the different ontologies. This task is referred to as **ontology matching** (Euzenat and Shvaiko 2007). The set of correspondences is called **alignment**. If the correspondence is directed, that is entities from one ontology map to others in another ontology (but not necessarily the reverse), the task is called **ontology mapping** (Kalfoglou and Schorlemmer 2003).

The main aim of ontology matching and mapping is to enable interoperability between systems using different knowledge models. Nonetheless, large-scale ontologies such as WordNet and Cyc have also been mapped (Medelyan and Legg 2008). Even semi-structured resources such as Wikipedia[17], whose semantics is only partially defined (Hovy, Navigli, and Ponzetto 2013), have been mapped to a lexical ontology such as WordNet, both when considering the category taxonomy of the Web encyclopedia (Ponzetto and Navigli 2009) and the graph structure induced by the hyperlinks within the pages (Navigli and Ponzetto 2012; Pilehvar and Navigli 2014).

Given the growing number of methods for ontology matching and mapping, an international competition called the **Ontology Alignment Evaluation Initiative (OAEI)**[18] is held every year at the Ontology Matching workshop held jointly with the International Semantic Web Conference.

Finally, similarly to what happens with schema integration in databases, different ontologies can be merged into a new ontology, a task referred to as **ontology merging**. An example of ontology merging was provided by integrating large-scale ontologies such as SENSUS, Cyc and Mikrokosmos (Hovy 1998).

20.7 Interfaces

Several interfaces to build and engineer ontologies have been proposed over the years. The importance of these tools lies in their ability to visually assist the ontology engineer in the creation, integration and maintenance phases. Among these tools, we mention:

- **OntoLingua**[19] (Farquhar, Fikes, and Rice 1997) – a Web distributed collaborative environment designed for viewing, creating and editing ontologies.
- **Protégé**[20] (Gennari et al. 2003) – a popular open source ontology editor written in Java with a large library of plugins for many applications,

including bioinformatics, Natural Language Processing, software engineering and validation.

- **OntoGen**[21] (Fortuna, Grobelnik, and Mladenic 2007) – a semi-automatic and data-driven editor for the creation and modification of ontologies.
- **Hozo**[22] (Kozaki et al. 2002) – an ontology editor based on a sophisticated ontological theory of roles.
- **WebODE**[23] (Corcho et al. 2002) – a Web tool for editing and modelling ontologies based on the METHONTOLOGY building approach.
- **SWOOP**[24] (Kalyanpur et al. 2006) – a Web tool aimed at fast and easy browsing and editing of ontologies, with support to collaborative annotation and versioning.
- **NeOn**[25] (Suarez-Figueroa and Gomez-Perez 2009) – a platform for browsing and manipulating ontologies, with a variety of plugins for annotation, development, reuse, acquisition, etc.
- **Altova SemanticWorks**[26] – a graphical environment for the visual development of ontologies.

20.8 Ontology Languages

Now that we know how to build or learn an ontology, what language are we supposed to use to encode it? And with what expressive power? This choice is crucial for enabling interoperability, semantic processing and reasoning, as languages that are too informal (e.g. just human-readable) or too expressive (e.g. first-order logic) might reduce the impact of ontological knowledge on intelligent systems.

Ontology languages are typically declarative and are commonly based either on first-order logic or on a fragment of it such as description logic. These include:

- **Knowledge Interchange Format (KIF)**: a knowledge representation language designed for exchange of knowledge between systems. It is based on LISP and first-order predicate logic.
- **Frame Logic (F-Logic)**: a declarative logic-based language designed to combine the advantages of ontological modelling with frame-based languages.
- **Common Logic**: a family of logic-based languages aimed at standardising the representation of syntax and semantics. Common Logic languages support first-order predicate logic, so they can be used to standardise first-order formulas.
- **CycL**: a declarative representation language based on first-order logic, with the addition of modal operators and higher-order quantification. It is used to represent the Cyc ontology.
- **Description Logics (DLs)**: a family of formal knowledge representation languages whose expressive power is between that of propositional logic and first-order predicate logic. A Description Logic (DL) models concepts and individuals, together with their relationships. The basic block of a DL is the **axiom**, that is a logical statement relating concepts and/or properties. Description logics distinguish between the so-called **TBox** (terminological box) and the **ABox** (assertional box). The former contains sentences describing relations between concepts, whereas the ABox

contains ground statements about individuals (e.g. relations between individuals and concepts).

- **Resource Description Framework (RDF)**[27]: a lightweight framework from W3C for the conceptual modelling of information identified by Web resources. The aim of RDF is to implement the vision of the Semantic Web in which Web resources are easily understood by machines thanks to semantic annotations. RDF provides a data model whose statements are triples of the form (*subject*, *property*, *object*), that can be written in XML format. The data model can be viewed as a graph, an example of which is shown in Figure 6 (strings are drawn as rectangles and URIs as ellipses). RDF triples in the graph are pairs of nodes (subject and object) connected by an edge (property). However, RDF concerns the ground level of an ontology, i.e. instances. To cope with concepts and relations W3C introduced a second language, called **RDF Schema (RDFS)**[28]. RDFS provides the syntax to define classes (i.e. concepts) and properties (i.e. relations), including a built-in *is-a* relationship. Recently, an RDF model for representing lexicalised ontologies has been put forward, called *lemon* (Lexicon Model for Ontology) (McCrae et al. 2012).[29] Large lexicalised ontologies such as BabelNet are now available in RDF-lemon format (Ehrmann et al. 2014).[30] The network of lexicalised resources represented in RDF and, in most cases, in RDF-lemon, is referred to as the Linguistic Linked Data cloud.[31]
- **Web Ontology Language (OWL)**[32], a family of knowledge representation languages for authoring ontologies endorsed by W3C. OWL builds upon RDF and RDFS and overcomes their limitations in terms of expressive

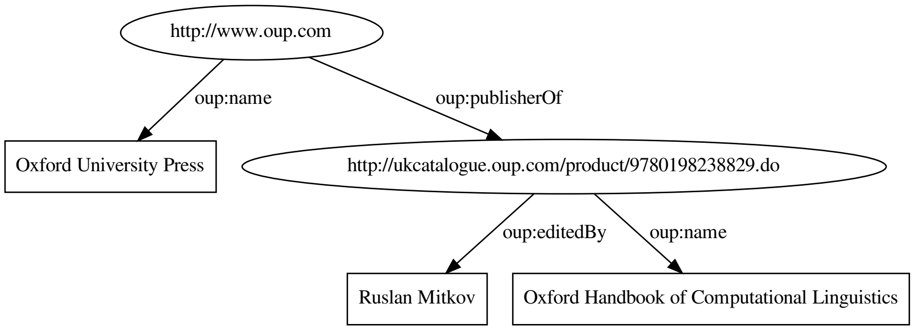

Figure 6
An example of an RDF graph.

power. OWL allows users to place restrictions on the cardinality of a property, to create new classes as union of other classes, to specify that two classes are disjoint (e.g. *plant* vs. *animal*), etc. There are three variants of OWL: a fully-expressive version (OWL Full), a computationally efficient version with the expressive power of Description Logics (OWL DL) and an easy-to-implement low-expressivity version (OWL Lite). Given that OWL is a standard for expressing ontologies, and thanks to the availability of several ontology editing tools (Mizoguchi and Kozaki 2009), it is among the most widespread ontology languages. However, many, who prefer more lightweight modelling, restrict to RDF(S).

Some of the above-mentioned languages are used in the Semantic Web layer cake (see Figure 3): XML is used to express the syntax of an ontology language, RDF for modelling instances, RDFS for encoding taxonomies (concepts and relations) and OWL for writing a full ontology.

20.9 Evaluation

Similarly to what happens with clustering techniques, evaluating an ontology is a key task that is difficult even for humans (also see Chapter 15). Indeed, it is very hard to find an objective way of assessing ontologies. One reason is that different ontologies might model the domain of interest equally well. Nonetheless, various different criteria have been proposed in the literature to assess the quality of an ontology. We can identify four main approaches to ontology evaluation (Brank, Grobelnik, and Mladenic 2005):

- **Human-based evaluation** using predefined criteria (Fox et al. 1998; Uschold and Jasper 1999; Burton-Jones et al. 2005; Gangemi et al. 2006; Obrst et al. 2007) or classifications (Hovy 2002). These include: *accuracy* (how close is the ontology model to the real world?), *adaptability* (how easily can the ontology be adapted/tailored to tasks, needs, etc.?), *clarity* (does the ontology encode the semantics of terms (concepts) in a way that is easy to understand?), *completeness* (how much of the domain does the ontology cover?), *conciseness* (how redundant is the ontology?), *efficiency* (how easily can the ontology be processed by reasoners and other intelligent systems?), *consistency* (are there contradictions in the ontology?). A thorough way of analysing ontologies is through OntoClean (Guarino and Welty 2002), a formal methodology for the analysis of ontologies based on conceptual properties that are independent of the domain. Another way of manually validating ontologies is through automatically-produced human-readable forms, obtained by combining textual definitions of the concepts linked through ontological relations (Navigli et al. 2004).

- **Comparison to a gold standard**: this kind of evaluation aims to compare the lexical and semantic structure of one or more ontologies with a humanly-created gold-standard ontology (see, e.g. Maedche and Staab (2002) and Kozareva and Hovy (2010b)). This approach has the advantage of performing one or more quantitative assessments of the ontologies of interest. However, it is not guaranteed that an ontology differing markedly from the gold standard is necessarily of low quality.
- **Task-based evaluation**, where the ontologies are plugged into an application and the output of the latter is evaluated in order to assess former. This evaluation approach has the advantage of avoiding the burden of evaluating a difficult artifact such as an ontology and indirectly assessing it on the basis of the performance increase or decrease produced by its use in an application (see, e.g. (Porzel and Malaka 2004)).
- **Data-driven evaluation**: this approach consists of using domain corpora (or other domain data) to assess the quality of one or more ontologies. An example of data-driven evaluation consists of automatically extracting terms from the corpus and then counting the number of terms extracted that are also contained in the ontology under evaluation (Brewster et al. 2004).

20.10 Applications

Ontologies are knowledge models, thus all the applications in need of structured knowledge can potentially benefit from their use. In this section we discuss popular, as well as potential, applications of ontologies, namely: the Semantic Web, Word Sense Disambiguation, automated reasoning, Question Answering, Semantic Information Retrieval, content-based Social Network Analysis and Machine Translation.

20.10.1 Semantic Web. In a sense, we could say that ontologies are the building blocks of the Semantic Web (see (Horrocks 2008) for a survey). In the Semantic Web vision, Web pages are semantically annotated with concepts, so as to provide an explicit meaning to be processed automatically. This ambitious vision can be implemented only if some kind of semantic "glue" is made available, i.e. if one or more ontologies are produced for each and every domain. As a result, applications such as semantic information retrieval and automatic reasoning, but also information sharing, question answering and content-based social network analysis, would be made possible.

20.10.2 Word Sense Disambiguation. Lexical ontologies, such as WordNet and BabelNet (cf. Section 20.4), have been shown to benefit Word Sense Disambiguation (WSD), the task of automatically associating meaning with words occurring in context (Navigli 2009, 2012) (see also Chapters 5 and 25). WSD systems exploiting ontological knowledge are called **knowledge-based**. It has been reported that knowledge-based systems perform as well as the best supervised systems on open texts (Ponzetto and Navigli 2010) and even outperform the best supervised systems on specific domains (Agirre and Soroa 2009; Ponzetto and Navigli 2010).

20.10.3 Automated reasoning. Automated reasoning is a subfield of Artificial Intelligence whose aim is to produce software systems that reason automatically. For instance, given the facts "Mario is Italian" and "Italians were born in Italy" we can infer that "Mario was born in Italy". Ontologies play a key role here, as they contain the knowledge needed to apply reasoning algorithms and thus infer new knowledge. In order to enable automated reasoning, ontologies need to be richly axiomatised and to avoid ambiguity as much as possible. Also, the ontology language chosen to encode the ontology (cf. Section 20.8) can impact heavily on decidability in reasoning. Popular software includes FaCT++[33] – an OWL-DL reasoner, Jena[34] – a Java framework that includes reasoning modules, PowerLoom [35] – a natural deduction inference engine based on a KIF variant, and Pellet[36] – a Java DL reasoner.

20.10.4 Question Answering. Another useful task in which ontologies have proven useful is Question Answering (QA, see also Chapter 36). QA aims at returning text snippets which provide an answer to a query expressed in natural language. Ontologies can be used to retrieve answer snippets that provide a reply to a target question but do not use the same words contained in the question (Mann 2002). For instance, given the question "Who is the current Bishop of Rome?", the system should be able to retrieve the answer "Benedict XVI" from the sentence "The current Pope is Benedict XVI".

Ontologies such as WordNet can be used in all three steps of a QA system (Paşca and Harabagiu 2001), namely: question processing (in determining the type and meaning of a question), passage retrieval (in formulating the most appropriate queries for identifying suitable passages) and answer extraction (identifying the portion of text which contains the answer). A well-known example of an ontology-based QA system is FALCON (Harabagiu et al. 2000).

20.10.5 Semantic Information Retrieval. A key problem in computer science is how to retrieve the desired information from large collections of documents such as the Web, a task referred to as Information Retrieval (see also Chapter 34). However, information is written in natural language, which is often ambiguous. An ideal information retrieval system should be able to effectively discard information containing the query words but concerning different senses (**polysemy**) and retrieve information satisfying the user needs, but expressed with different words (**synonymy**).

Ontologies can be used to perform semantically-informed Information Retrieval (see also Chapter 34). Over the years, different methods have been proposed (Krovetz and Croft 1992; Voorhees 1993; Mandala, Tokunaga, and Tanaka 1998; Gonzalo, Penas, and Verdejo 1999, inter alia). However, contrasting results have been reported on the benefits of these techniques: given that Word Sense Disambiguation (see also Chapter 25) is involved, it has been shown that the semantic annotation step has to be very accurate to benefit Information Retrieval (Sanderson 1994) – a result that was later debated (Gonzalo, Penas, and Verdejo 1999; Stokoe, Oakes, and Tait 2003). Finally, interesting results have been reported on ontology-based query expansion when expanding queries with words from textual definitions of query concepts in the lexical ontology (Navigli and Velardi 2005).

20.10.6 Content-based Social Network Analysis. Social network analysis (SNA) is the field studying the relationships between people, organisations, animals, etc. The study is conducted by means of methods from network theory, where a network consists of nodes (the entities of interest) and edges (i.e. links or connections between the entities). Ontologies can be of help to SNA for many reasons, the most immediate one coming from their very nature: they encode a network of relations between entities, thus they can be used to encode knowledge about social connections. Furthermore, ontologies can be used to discover or infer new knowledge about social networks, e.g. when dealing with terrorism data (Wennerberg 2005), or to semantically analyse the communicative content of the social network (Velardi et al. 2008).

20.10.7 Machine Translation. Machine Translation (MT, see also Chapter 32) is a long-standing topic in computational linguistics. In the last two decades, statistical machine translation has been shown to provide the best results. However, these methods lack a real understanding of the semantics of text. While we are far from performing semantically-informed MT, approaches have been proposed that use an interlingua as an intermediate representation of meaning (Nirenburg, Raskin, and Tucker 1986), automatically translate terminology by means of ontology learning (Navigli, Velardi, and Gangemi 2003), as well as iteratively improve the performance of MT by means of a multilingual ontology (Knoth et al. 2010).

20.11 Conclusions

Ontologies are semantic data structures that provide an explicit modelling for a portion of the real world. As such, they help scientists, linguistics and philosphers to crystallise knowledge. Further, given that knowledge is expressed through language, most ontologies are lexicalised, ranging from domain-specific to general-purpose ones. As a consequence, all language-based areas of computer science can be semantically enabled, including text annotation, disambiguation, processing, analysis, translation and retrieval.

We believe that the next challenge is to make medium-sized and large-scale ontologies available for many domains, provide mappings for them so as to enable interoperability, and inject semantics into current off-line and on-line applications, with the ambitious objective of putting into practice the exciting vision of the Semantic Web.

Further reading and relevant resources

A number of introductions to ontologies can be found on-line[37] as well as entire books devoted to the topic, some focusing more on Semantic Web aspects (Staab and Studer 2009), others more concerned with a computational linguistics perspective on the topic (Huang et al. 2010).

Many ontology repositories are accessible on-line, such as the Semantic Web repository[38] – which contains a list of basic upper and domain ontologies, the TONES repository[39] – a central ontology deposit created in the context of an EU FET project, and the Swoogle ontology search facility[40] – that stores and indexes "Semantic Web documents", i.e. documents written in RDF crawled from the Web. The Sweet Compendium of Ontology Building Tools[41] provides an up-to-date list with dozens of links to ontology building and learning tools. An "intrepid guide to ontologies"[42] is also available from the same author, Mike Bergman. The Global WordNet Association (GWA)[43] fosters the discussion, sharing and interconnection of wordnets for all languages in the world. The recent LIDER project[44] has been fostering the creation of a Linguistic Linked Data cloud.

Journals dealing with various different aspects of ontologies include: Computational Linguistics (MIT Press), Natural Language Engineering (Cambridge University Press), IEEE Transactions on Knowledge and Data Engineering (IEEE Press), Data & Knowledge Engineering (Elsevier), Journal of Web Semantics (Elsevier), Artificial Intelligence (Elsevier), Journal of Artificial Intelligence Reseach (AAAI Press), and many others. Conferences include: ACL, IJCAI, AAAI, EMNLP, EACL, ISWC, ESWC, EKAW, FOIS, LREC, GWC. Many workshops have been organised on the topic of ontologies, including the following series: Ontology Learning and Population (OLP), Linked Data on the Web (LDOW), Ontology Matching (OM), Semeval (formerly Senseval) on semantic evaluation, Ontologies and Semantic Web for E-Learning (SWEL), Vocabularies, Ontologies and Rules for The Enterprise (VORTE).

Notes

[1] http://www.w3.org/TR/REC-xml/

[2] http://www.w3.org/TR/rdf-sparql-query/

[3] http://linkeddata.org/

[4] http://ontologydesignpatterns.org

[5] http://www.ontologyportal.org

[6] http://wordnet.princeton.edu

[7] http://babelnet.org

[8] http://wibitaxonomy.org

[9] http://www.cyc.com

[10] http://www.mpi-inf.mpg.de/yago-naga/yago/

[11] http://dbpedia.org

[12] http://omega.isi.edu

[13] http://www.nlm.nih.gov/research/umls/

[14] http://www.ihtsdo.org/snomed-ct

[15] http://www.geneontology.org

[16] http://pir.georgetown.edu/pro/

[17] http://www.wikipedia.org

[18] http://oaei.ontologymatching.org/

[19] http://ksl.stanford.edu/software/ontolingua/

[20] http://protege.stanford.edu/

[21] http://ontogen.ijs.si/

[22] http://www.hozo.jp/

[23] http://webode.dia.fi.upm.es

[24] http://code.google.com/p/swoop/

[25] http://www.neon-project.org

[26] http://www.altova.com/semanticworks.html

[27] http://www.w3.org/RDF/

[28] http://www.w3.org/TR/rdf-schema/

[29] http://lemon-model.net

[30] http://babelnet.org/rdf

[31] http://linghub.lider-project.eu/llod-cloud

[32] http://www.w3.org/TR/owl-features/

[33] http://owl.man.ac.uk/factplusplus/

[34] http://openjena.org

[35] http://www.isi.edu/isd/LOOM/PowerLoom/

[36] http://clarkparsia.com/pellet/

[37] E.g., http://www.mt-archive.info/AMTA-2006-Hovy.pdf

[38] http://semanticweb.org/wiki/Ontology

[39] http://owl.cs.manchester.ac.uk/repository/

[40] http://swoogle.umbc.edu

[41] http://www.mkbergman.com/862/the-sweet-compendium-of-ontology-building-tools/

[42] http://www.mkbergman.com/374/an-intrepid-guide-to-ontologies/

[43] http://www.globalwordnet.org

[44] http://lider-project.eu

Acknowledgments

This chapter is dedicated to the memory of Gianni Chilanti, a friend like no other. Special thanks go to Ed Hovy and Jim McManus for their very useful comments and suggestions on the chapter. Figure 1 first appeared in (Navigli and Velardi 2008), Figures 4 and 5 first appeared in (Navigli 2009), Figure 3 is distributed in the public domain (see http://en.wikipedia.org/wiki/File:Semantic-web-stack.png). The author gratefully acknowledges the support of the ERC Starting Grant MultiJEDI No. 259234 and the EU LIDER project No. 610782.

References

[Agirre and Soroa2009]Agirre, Eneko and Aitor Soroa. 2009. Personalizing PageRank for Word Sense Disambiguation. In *Proceedings of the 12th conference of the European chapter of the Association for Computational Linguistics (EACL-2009)*, pages 33–41, Athens, Greece.

[Auer et al.2007]Auer, Sören, Christian Bizer, Georgi Kobilarov, Jens Lehmann, and Zachary Ives. 2007. DBpedia: A nucleus for a web of open data. In *Proceedings of 6th International Semantic Web Conference (ISWC)*, pages 11–15, Busan, Korea. Springer.

[Bond and Paik2012]Bond, Francis and Kyonghee Paik. 2012. A survey of wordnets and their licenses. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*, page 64âĂŞ71, Matsue.

[Borgo, Guarino, and Masolo1996]Borgo, Stefano, Nicola Guarino, and Claudio Masolo. 1996. A pointless theory of space based on strong connection and congruence. In *Proceedings of Principles of Knowledge Representation and Reasoning (KR96)*, pages 220–229, Cambridge, Massachusetts, USA. Morgan Kaufmann.

[Brank, Grobelnik, and Mladenic2005]Brank, Janez, Marko Grobelnik, and Dunja Mladenic. 2005. A survey of ontology evaluation techniques. In *Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005)*, pages 166–170, Ljubljana, Slovenia.

[Brewster et al.2004]Brewster, Christopher, Harith Alani, Srinandan Dasmahapatra, and Yorick Wilks. 2004. Data-driven ontology evaluation. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2004)*, pages 164–168, Lisbon, Portugal. European Language Resources Association.

[Buitelaar, Olejnik, and Sintek2004]Buitelaar, Paul, Daniel Olejnik, and Michael Sintek. 2004. A Protege plug-in for ontology extraction from text based on linguistic analysis. In *Proceedings of the 1st European Semantic Web Symposium (ESWS)*, pages 31–44, Heraklion, Greece.

[Burton-Jones et al.2005]Burton-Jones, Andrew, Veda C. Storey, Vijayan Sugumaran, and Punit Ahluwalia. 2005. A semiotic metrics suite for assessing the quality of ontologies. *Data & Knowledge Engineering*, 55(1):84–102.

[Carlson et al.2010]Carlson, Andrew, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka Jr., and Tom M. Mitchell. 2010. Toward an architecture for never-ending language learning. In *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence (AAAI 2010)*, pages 1306–1313, Atlanta, Georgia.

[Cimiano, Hotho, and Staab2005]Cimiano, Philipp, Andreas Hotho, and Steffen Staab. 2005. Learning concept hierarchies from text corpora using formal concept analysis. *Journal of Artificial Intelligence Research*, 24(1):305–339.

[Cimiano and Völker2005]Cimiano, Philipp and Johanna Völker. 2005. Text2Onto - a framework for ontology learning and data-driven change discovery. In Andres Montoyo, Rafael Munoz, and Elisabeth Metais, editors, *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems (NLDB)*, volume 3513 of *Lecture Notes in Computer Science*, pages 227–238, Alicante, Spain. Springer.

[Corcho et al.2002]Corcho, Oscar, Mariano FernÃąndez-LÃşpez, AsunciÃşn GÃşmez-PÃľrez, and Oscar Vicente. 2002. WebODE: An integrated workbench for ontology representation, reasoning, and exchange. In *Proceedings of the Conference on Knowledge Acquisition, Modeling and Management (EKAW)*, pages 138–153. Springer.

[Crofts et al.2010]Crofts, Nick, Martin Doerr, Tony Gill, Stephen Stead, and Matthew Stiff. 2010. *Definition of the CIDOC Conceptual Reference Model*. ICOM/CIDOC CRM Special Interest Group.

[De Nicola, Missikoff, and Navigli2009]De Nicola, Antonio, Michele Missikoff, and Roberto Navigli. 2009. A software engineering approach to ontology building. *Information Systems*, 34(2):258–275.

[Ehrmann et al.2014]Ehrmann, Maud, Francesco Cecconi, Daniele Vannella, John Philip McCrae, Philipp Cimiano, and Roberto Navigli. 2014. Representing multilingual data as linked data: the case of babelnet 2.0. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), Reykjavik, Iceland, May 26-31, 2014.*, pages 401–408.

[Etzioni et al.2004]Etzioni, Oren, Michael Cafarella, Doug Downey, Stanley Kok, Ana-Maria Popescu, Tal Shaked, Stephen Soderland, Daniel S. Weld, and Alexander Yates. 2004. Web-scale information extraction in KnowItAll. In *Proceedings of the* 13^{th} *International Conference on World Wide Web (WWW)*, pages 100–110, New York, NY, USA.

[Euzenat and Shvaiko2007]Euzenat, Jérôme and Pavel Shvaiko. 2007. *Ontology matching*. Springer-Verlag, Heidelberg, Germany.

[Fader, Soderland, and Etzioni2011]Fader, Anthony, Stephen Soderland, and Oren Etzioni. 2011. Identifying relations for open information extraction. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, EMNLP 2011*, pages 1535–1545, Edinburgh, UK.

[Farquhar, Fikes, and Rice1997]Farquhar, Adam, Richard Fikes, and James Rice. 1997. The Ontolingua Server: a tool for collaborative ontology construction. *International Journal of*

Human-Computer Studies, 46(6):707–727.

[Fellbaum1998]Fellbaum, Christiane, editor. 1998. *WordNet: An Electronic Database*. MIT Press, Cambridge, MA, USA.

[Fernandez-Lopez, Gomez-Perez, and Juristo1997]Fernandez-Lopez, Mariano, Asuncion Gomez-Perez, and Natalia Juristo. 1997. Methontology: from ontological art towards ontological engineering. In *Proceedings of the AAAI97 Spring Symposium*, pages 33–40, Stanford, USA.

[Ferragina and Scaiella2010]Ferragina, Paolo and Ugo Scaiella. 2010. TAGME: On-the-fly Annotation of Short Text Fragments (by Wikipedia Entities). In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 1625–1628, New York, NY, USA. ACM.

[Flati et al.2014]Flati, Tiziano, Daniele Vannella, Tommaso Pasini, and Roberto Navigli. 2014. Two Is Bigger (and Better) Than One: the Wikipedia Bitaxonomy Project. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*, pages 945–955, Baltimore, Maryland. Association for Computational Linguistics.

[Fortuna, Grobelnik, and Mladenic2007]Fortuna, Blaz, Marko Grobelnik, and Dunja Mladenic. 2007. OntoGen: Semi-automatic ontology editor. In *Proceedings of the 12th International Conference on Human-Computer Interaction (HCI)*, pages 309–318, Beijing, China.

[Fox et al.1998]Fox, Mark S., Mihai Barbuceanu, Michael Gruninger, and Jinxin Lin. 1998. An organization ontology for enterprise modelling. *Simulating organizations: Computational models of institutions and groups*, pages 123âĂŞ134.

[Gangemi et al.2006]Gangemi, Aldo, Carola Catenacci, Massimiliano Ciaramita, and Jos Lehmann. 2006. Modelling ontology evaluation and validation. In *Proceedings of the 3rd European Semantic Web Conference (ESWC 2006)*, pages 140–154, Budva, Montenegro.

[Gangemi and Presutti2009]Gangemi, Aldo and Valentina Presutti, 2009. *Handbook of Ontologies (2nd edition)*, chapter Ontology Design Patterns, pages 221–243. Springer, Berlin, Germany.

[Gennari et al.2003]Gennari, John H., Mark A. Musen, Ray W. Fergerson, William E. Grosso, Monica Crubézy, Henrik Eriksson, Natalya F. Noy, and Samson W. Tu. 2003. The evolution of Protégé: an environment for knowledge-based systems development. *International Journal of Human-Computer Studies*, 58:89–123, January.

[Gonzalo, Penas, and Verdejo1999]Gonzalo, Julio, Anselmo Penas, and Felisa Verdejo. 1999. Lexical ambiguity and Information Retrieval revisited. In *Proceedings of the 1999 Conference on Empirical Methods in NLP and Very Large Corpora (EMNLP/VLC)*, pages 195–202, College Park, MD, USA.

[Gruber1993]Gruber, Thomas R. 1993. Toward principles for the design of ontologies used for knowledge sharing. In *International Workshop on Formal Ontology*, pages 907–928, Padova, Italy.

[Guarino and Welty2002]Guarino, Nicola and Chris Welty. 2002. Evaluating ontological decisions with OntoClean. *Communications of the ACM*, 45(2):61–65.

[Harabagiu et al.2000]Harabagiu, Sanda, Dan Moldovan, Marius Pasca, Rada Mihalcea, Mihai Surdeanu, Razvan Bunescu, Roxana Girju, Vasile Rus, and Paul Morarescu. 2000. Falcon: Boosting knowledge for answer engine. In *Proceedings of the* 9^{th} *Text REtrieval Conference (TREC-9)*, pages 479–488, Gaithersburg, MD.

[Hasegawa, Sekine, and Grishman2004]Hasegawa, Takaaki, Satoshi Sekine, and Ralph Grishman. 2004. Discovering relations among named entities from large corpora. In *Proceedings of the 42nd Annual Meeting of Association of Computational Linguistics (ACL 04)*, pages 415–422, Barcelona.

[Hearst1992]Hearst, Marti. 1992. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the* 14^{th} *International Conference on Computational Linguistics (COLING)*, pages 539–545, Nantes, France.

[Hirst2009]Hirst, Graeme, 2009. *Handbook on Ontologies (second edition)*, chapter Ontology and the lexicon, pages 269–292. Springer, Berlin, Germany.

[Hoffart et al.2013]Hoffart, Johannes, Fabian M. Suchanek, Klaus Berberich, and Gerhard Weikum. 2013. YAGO2: A spatially and temporally enhanced knowledge base from wikipedia. *Artif. Intell.*, 194:28–61.

[Horrocks2008]Horrocks, Ian. 2008. Ontologies and the semantic web. *Communications of the ACM*, 51:58–67, December.

[Hovy1998]Hovy, Eduard. 1998. Combining and Standardizing Large-Scale, Practical Ontologies for Machine Translation and Other Uses. In *Proceedings of the Language Resource and Evaluation Conference (LREC)*, Granada, Spain.

[Hovy2002]Hovy, Eduard H. 2002. Comparing Sets of Semantic Relations in Ontologies. In R. Green, C. A. Bean, and Myaeng S. H. (eds), editors, *The Semantics of Relationships: An Interdisciplinary Perspective*. Dordrecht: Kluwer, pages 91–110.

[Hovy, Navigli, and Ponzetto2013]Hovy, Eduard H., Roberto Navigli, and Simone Paolo Ponzetto. 2013. Collaboratively built semi-structured content and Artificial Intelligence: The story so far. *Artificial Intelligence*, 194:2–27.

[Huang et al.2010]Huang, Chu-ren, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prevot, editors. 2010. *Ontology and the Lexicon*. Cambridge University Press.

[Jain and Pantel2010]Jain, Alpa and Patrick Pantel. 2010. FactRank: Random walks on a Web of facts. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, pages 501–509, Beijing, China, August.

[Jarmasz and Szpakowicz2003]Jarmasz, Mario and Stan Szpakowicz. 2003. Roget's thesaurus and semantic similarity. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP-03)*, pages 212–219, Borovets, Bulgaria.

[Jiang and Conrath1997]Jiang, Jay J. and David W. Conrath. 1997. Semantic similarity based on corpus statistics and lexical taxonomy. In *Proceedings of* 10^{th} *International Conference on Research in Computational Linguistics*, pages 19–33, Taiwan.

[Kalfoglou and Schorlemmer2003]Kalfoglou, Yannis and Marco Schorlemmer. 2003. Ontology mapping: the state of the art. *Knowledge Engineering Review*, 18:1–31, January.

[Kalyanpur et al.2006]Kalyanpur, Aditya, Bijan Parsia, Evren Sirin, Bernardo C. Grau, and James Hendler. 2006. Swoop: A web ontology editing browser. *Journal of Web Semantics*, 4(2):144–153.

[Kifer2008]Kifer, Michael. 2008. Rule interchange format: The framework. In *Proceedings of the 2nd International Conference on Web Reasoning and Rule Systems*, RR '08, pages 1–11, Berlin, Heidelberg. Springer-Verlag.

[Knoth et al.2010]Knoth, Petr, Trevor Collins, Elsa Sklavounou, and Zdenek Zdrahal. 2010. Facilitating cross-language retrieval and machine translation by multilingual domain ontologies. In *Proceedings of the 7th Conference on Language Resources and Evaluation (LREC)*, Valletta, Malta.

[Kozaki et al.2002]Kozaki, Kouji, Yoshinobu Kitamura, Mitsuru Ikeda, and Riichiro Mizoguchi. 2002. Hozo: An environment for building/using ontologies based on a fundamental consideration of "role" and "relationship". In *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW)*, EKAW '02, pages 213–218, London, UK. Springer-Verlag.

[Kozareva and Hovy2010a]Kozareva, Zornitsa and Eduard Hovy. 2010a. Learning arguments and supertypes of semantic relations using recursive patterns. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1482–1491, Uppsala, Sweden, July.

[Kozareva and Hovy2010b]Kozareva, Zornitsa and Eduard Hovy. 2010b. A semi-supervised method to learn and construct taxonomies using the web. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1110–1118, Cambridge, MA, October.

[Krovetz and Croft1992]Krovetz, Robert and William B. Croft. 1992. Lexical ambiguity and information retrieval. *ACM Transactions on Information Systems*, 10(2):115–141.

[Leacock and Chodorow1998]Leacock, Claudia and Martin Chodorow. 1998. Combining local context and WordNet similarity for word sense identification. In Christiane Fellbaum, editor, *WordNet: An electronic lexical database*. MIT Press, Cambridge, MA, USA, pages 265–283.

[Lenat1995]Lenat, Douglas B. 1995. Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11).

[Maedche and Staab2000]Maedche, Alexander and Steffen Staab. 2000. Discovering conceptual relations from text. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI)*, pages 321–325, Berlin, Germany.

[Maedche and Staab2002]Maedche, Alexander and Steffen Staab. 2002. Measuring similarity between ontologies. In *Proceedings of the European Conference on Knowledge Acquisition and Management (EKAW)*, pages 251–263, Madrid, Spain.

[Maedche and Volz2001]Maedche, Alexander and Raphael Volz. 2001. The Text-To-Onto Ontology Extraction and Maintenance System. In *Workshop on Integrating Data Mining and Knowledge Management co-located with the 1st International Conference on Data Mining*, pages 478–481, San Jose, California, USA, 11.

[Magnini and Cavaglià2000]Magnini, Bernardo and Gabriela Cavaglià. 2000. Integrating subject field codes into wordnet. In *Proceedings of the 2^{nd} Conference on Language Resources and Evaluation (LREC)*, pages 1413–1418, Athens, Greece.

[Mandala, Tokunaga, and Tanaka1998]Mandala, Rila, Takenobu Tokunaga, and Hozumi Tanaka. 1998. The use of WordNet in Information Retrieval. In *Proceedings of the COLING-ACL workshop on Usage of Wordnet in Natural Language Processing*, pages 31–37, Montreal, Canada.

[Mann2002]Mann, Gideon S. 2002. Fine-grained proper noun ontologies for question answering. In *COLING-02 on SEMANET: building and using semantic networks - Volume 11*, pages 1–7, Taipei, Taiwan.

[Masolo et al.2003]Masolo, Claudio, Stefano Borgo, Aldo Gangemi, Nicola Guarino, and Alessandro Oltramari. 2003. Ontology library. In *WonderWeb Deliverable D18*.

[McCrae et al.2012]McCrae, John, Guadalupe Aguado de Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asuncion Gómez-Pérez, Jorge Gracia, Laura Hollink, Elena Montiel-Ponsoda, Dennis Spohr, and Tobias Wunner. 2012. Interchanging lexical resources on the semantic web. *Language Resources and Evaluation*, 46(4):701–719.

[McCray and Nelson1995]McCray, Alexa T. and Stuart J. Nelson. 1995. The representation of meaning in the umls. *Methods of Information in Medicine*, 34:193–201.

[Medelyan and Legg2008]Medelyan, Olena and Catherine Legg. 2008. Integrating Cyc and Wikipedia: Folksonomy meets rigorously defined common-sense. In *Proceedings of the WIKI-AI: Wikipedia and AI Workshop at the AAAI*, volume 8, pages 13–18.

[Mizoguchi and Kozaki2009]Mizoguchi, Riichiro and Kouji Kozaki, 2009. *Handbook on Ontologies (second edition)*, chapter Ontology Engineering Environments, pages 315–336. Springer, Berlin, Germany.

[Moro and Navigli2013]Moro, Andrea and Roberto Navigli. 2013. Integrating syntactic and semantic analysis into the open information extraction paradigm. In *Proceedings of the 23^{rd} International Joint Conference on Artificial Intelligence*, pages 2148–2154, Beijing, China.

[Moro, Raganato, and Navigli2014]Moro, Andrea, Alessandro Raganato, and Roberto Navigli. 2014. Entity Linking meets Word Sense Disambiguation: a Unified Approach. *Transactions of the Association for Computational Linguistics (TACL)*, 2:231–244.

[Natale et al.2006]Natale, Darren A., Cecilia N. Arighi, Winona Barker, Judith Blake, Ti-Cheng Chang, Zhangzhi Hu, Hongfang Liu, Barry Smith, and Cathy H. Wu. 2006. Framework for a protein ontology. In *Proceedings of the 1st international workshop on Text mining in bioinformatics*, TMBIO '06, pages 29–36, New York, NY, USA. ACM.

[Navigli2009]Navigli, Roberto. 2009. Word Sense Disambiguation: a survey. *ACM Computing Surveys*, 41(2):1–69.

[Navigli2012]Navigli, Roberto. 2012. A quick tour of word sense disambiguation, induction and related approaches. In *Proceedings of the 38th Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM)*, pages 115–129.

[Navigli and Lapata2010]Navigli, Roberto and Mirella Lapata. 2010. An experimental study of graph connectivity for unsupervised word sense disambiguation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):678–692.

[Navigli and Ponzetto2012]Navigli, Roberto and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.

[Navigli and Velardi2004]Navigli, Roberto and Paola Velardi. 2004. Learning domain ontologies from document warehouses and dedicated websites. *Computational Linguistics*, 30(2):151–179.

[Navigli and Velardi2005]Navigli, Roberto and Paola Velardi. 2005. Structural Semantic Interconnections: a knowledge-based approach to Word Sense Disambiguation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(7):1075–1088.

[Navigli and Velardi2008]Navigli, Roberto and Paola Velardi. 2008. From glossaries to ontologies: Extracting semantic structure from textual definitions. In Paul Buitelaar and Philipp Cimiano, editors, *Ontology Learning and Population: Bridging the Gap between Text and Knowledge*, pages 71–87, Amsterdam, The Netherlands. IOS Press.

[Navigli and Velardi2010]Navigli, Roberto and Paola Velardi. 2010. Learning word-class lattices for definition and hypernym extraction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, pages 1318–1327, Uppsala, Sweden, July.

[Navigli et al.2004]Navigli, Roberto, Paola Velardi, Alessandro Cucchiarelli, and Francesca Neri. 2004. Quantitative and qualitative evaluation of the OntoLearn ontology learning system. In *Proceedings of* 20^{th} *International Conference on Computational Linguistics (COLING 2004)*, pages

1043–1050, Geneva, Switzerland.

[Navigli, Velardi, and Gangemi2003]Navigli, Roberto, Paola Velardi, and Aldo Gangemi. 2003. Ontology learning and its application to automated terminology translation. *IEEE Intelligent Systems*, 18(1):22–31.

[Nirenburg, Raskin, and Tucker1986]Nirenburg, Sergei, Victor Raskin, and Allen Tucker. 1986. On knowledge-based machine translation. In *Proceedings of the 11th Conference on Computational linguistics (COLING)*, pages 627–632, Morristown, NJ, USA. Association for Computational Linguistics.

[Obrst et al.2007]Obrst, Leo, Werner Ceusters, Inderjeet Mani, Steve Ray, and Barry Smith. 2007. The Evaluation of Ontologies: Toward Improved Semantic Interoperability. In *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences*. pages 139–158.

[Paşca and Harabagiu2001]Paşca, Marius and Sanda Harabagiu. 2001. The informative role of wordnet in open-domain question answering. In *Proceedings of the NAACL 2001 Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations*, pages 138–143, Carnegie Mellon University, Pittsburgh PA.

[Paşca et al.2006]Paşca, Marius, Dekang Lin, Jeffrey Bigham, Andrei Lifchits, and Alpa Jain. 2006. Names and similarities on the web: Fact extraction in the fast lane. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL)*, pages 809–816, Sydney, Australia, July.

[Pantel and Pennacchiotti2006]Pantel, Patrick and Marco Pennacchiotti. 2006. Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL)*, pages 113–120, Sydney, Australia.

[Park, Byrd, and Boguraev2002]Park, Youngja, Roy J Byrd, and Branimir K Boguraev. 2002. Automatic glossary extraction: beyond terminology identification. In *Proceedings of the 19th international conference on Computational linguistics (COLING)*, pages 1–7, Taipei, Taiwan.

[Pease, Niles, and Li2002]Pease, Adam, Ian Niles, and John Li. 2002. The Suggested Upper Merged Ontology: A large ontology for the semantic Web and its applications. In *Proceedings of AAAI-2002 Workshop on Ontologies and the Semantic Web*, Edmonton, Canada.

[Philpot, Hovy, and Pantel2005]Philpot, Andrew, Eduard Hovy, and Patrick Pantel. 2005. The Omega Ontology. In *Proceedings of IJCNLP workshop on Ontologies and Lexical Resources (OntoLex-05)*, pages 59–66, Jeju Island, South Korea.

[Pilehvar, Jurgens, and Navigli2013]Pilehvar, Mohammad Taher, David Jurgens, and Roberto Navigli. 2013. Align, Disambiguate and Walk: a Unified Approach for Measuring Semantic Similarity. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 1341–1351, Sofia, Bulgaria.

[Pilehvar and Navigli2014]Pilehvar, Mohammad Taher and Roberto Navigli. 2014. A robust approach to aligning heterogeneous lexical resources. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*, pages 468–478, Baltimore, Maryland. Association for Computational Linguistics.

[Ponzetto and Navigli2009]Ponzetto, Simone Paolo and Roberto Navigli. 2009. Large-scale taxonomy mapping for restructuring and integrating Wikipedia. In *Proceedings of the* 21^{st} *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2083–2088, Pasadena, California.

[Ponzetto and Navigli2010]Ponzetto, Simone Paolo and Roberto Navigli. 2010. Knowledge-rich Word Sense Disambiguation rivaling supervised system. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, pages 1522–1531, Uppsala, Sweden.

[Porzel and Malaka2004]Porzel, Robert and Rainer Malaka. 2004. A task-based approach for ontology evaluation. In *Proceedings of ECAI 2004 Workshop on Ontology Learning and Population*, Valencia, Spain, August.

[Rapp2003]Rapp, R. 2003. Word sense discovery based on sense descriptor dissimilarity. In *Proceedings of the Ninth Machine Translation Summit*, pages 315–322, New Orleans, USA.

[Reiter and Criscuolo1981]Reiter, Raymond and G. Criscuolo. 1981. On interacting defaults. In *Proceedings of the* 7^{th} *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 270–276.

[Sanderson1994]Sanderson, Mark. 1994. Word sense disambiguation and information retrieval. In *Proceedings of the Special Interest Group on Information Retrieval (SIGIR)*, pages 142–151, Dublin, Ireland.

[Seedorf, Informatik, and Mannheim2006]Seedorf, Stefan, Fzi Forschungszentrum Informatik, and UniversitÃd't Mannheim. 2006. Applications of ontologies in software engineering. In *Proceedings of 2nd International Workshop on Semantic Web Enabled Software Engineering (SWESE 2006), joint with the 5th International Semantic Web Conference (ISWC 2006)*.

[Sowa2000]Sowa, John F. 2000. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co., Pacific Grove, CA.

[Staab and Studer2009]Staab, Steffen and Rudi Studer, editors. 2009. *Handbook on Ontologies*. Springer, Berlin, 2. edition.

[Stokoe, Oakes, and Tait2003]Stokoe, Christopher, Michael J. Oakes, and John I. Tait. 2003. Word sense disambiguation in Information Retrieval revisited. In *Proceedings of the 26^{th} Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 159–166, Toronto, Canada.

[Suarez-Figueroa and Gomez-Perez2009]Suarez-Figueroa, Mari Carmen and Asuncion Gomez-Perez. 2009. Neon methodology for building ontology networks: a scenario-based methodology. In *Proceedings of the International Conference on Software, Services & Semantic Technologies (S3T 2009)*.

[Suchanek, Kasneci, and Weikum2007]Suchanek, Fabian M., Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web (WWW)*, pages 697–706, New York, NY, USA. ACM.

[Sure et al.2003]Sure, York, Steffen Staab, Rudi Studer, and Ontoprise Gmbh. 2003. On-To-Knowledge methodology (OTKM). In *Handbook on Ontologies, International Handbooks on Information Systems*, pages 117–132. Springer.

[The Gene Ontology Consortium2008]The Gene Ontology Consortium. 2008. The Gene Ontology project in 2008. *Nucl. Acids Res.*, 36(suppl_1):D440–444.

[Turney et al.2003]Turney, Peter D., Michael L. Littman, Jeffrey Bigham, and Victor Shnayder. 2003. Combining independent modules to solve multiple-choice synonym and analogy problems. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP-03)*, pages 482–489, Borovets, Bulgaria.

[Uschold and Jasper1999]Uschold, Michael and Rob Jasper. 1999. A framework for understanding and classifying ontology applications. In *Twelfth Workshop on Knowledge*

Acquisition Modeling and Management KAW 99.

[Velardi, Faralli, and Navigli2013]Velardi, Paola, Stefano Faralli, and Roberto Navigli. 2013. OntoLearn Reloaded: A Graph-Based Algorithm for Taxonomy Induction. *Computational Linguistics*, 39(3):665–707.

[Velardi et al.2008]Velardi, Paola, Roberto Navigli, Alessandro Cucchiarelli, and Fulvio D'Antonio. 2008. A new content-based model for social network analysis. In *Proceedings of the 2008 IEEE International Conference on Semantic Computing (ICSC 2008)*, pages 18–25, Washington, DC, USA. IEEE Computer Society.

[Velardi, Navigli, and D'Amadio2008]Velardi, Paola, Roberto Navigli, and Pierluigi D'Amadio. 2008. Mining the web to create specialized glossaries. *IEEE Intelligent Systems*, 23(5):18–25.

[Völker, Hitzler, and Cimiano2007]Völker, Johanna, Pascal Hitzler, and Philip Cimiano. 2007. Acquisition of OWL DL axioms from lexical resources. In *Proceedings of the 4th European Semantic Web Conference (ESWC 2007)*, pages 670–685, Innsbruck, Austria. Springer.

[Voorhees1993]Voorhees, Ellen M. 1993. Using wordnet to disambiguate word senses for text retrieval. In *Proceedings of the* 16^{th} *Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 171–180, Pittsburgh, PA, USA.

[Vossen1998]Vossen, Piek, editor. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Kluwer, Dordrecht, The Netherlands.

[Wennerberg2005]Wennerberg, Pinar O. 2005. Ontology based knowledge discovery in social networks. *JRC Joint Research Center*.