

## Question 2

Today's content on the web is mostly human readable and not machine readable. This it difficult for machines to independently, form a human intelligence, find and deduce information. Current systems can only to text searches, but cannot deduce meaning. Hence the development of the Semantic Web. "The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users". (Berners-Lee, Hendler, & Lassila, 2001) Examples of services the Semantic Web will enable are information brokers, search agents and information filters. (Decker et al., 2000)

Agents will be able to deduce the meaning of the data and infer new knowledge from what it already knows. Because the agents will be using the semantics embedded in the web pages, we will not need AI with a level of sophistication of a human being. (Berners-Lee et al., 2001)

The semantic information has to be encoded into web pages using a machine readable language for these agents to function because semantics of data is rarely explicit. (Heiler, 1995) With this in mind, there are a couple of languages that have come forward in which we can do this encoding. RDF and RDF SCHEMA were the initial focus of the but we found to be lacking in expressive power. The World Wide Web Consortium (W3C) created the Web Ontology Group to develop an Ontology language for the Semantic Web. The relationship between terms are to be provided by this ontology in a structured vocabulary. Making it easy for agent to interpret unambiguously. (Horrocks, Patel-Schneider, & van Harmelen, 2003) The language developed by the Web Ontology Group is called Web Ontology Language or OWL.

Using common languages, allows agents to search through different web pages to find relevant data. But using the same language, like OWL, does not guarantee that the meaning of particular terms are the same across different web pages or domains. We can guarantee *semantic interoperability* (meaning of the data) by using the same standard ontology. (van Diggelen & Dignum, 2007). In such a case the meaning of all terms and relationships are the same by definition. The definition of the ontology.

But there are very many websites and databases that do not use the same ontologies. Different domains use different ontologies because each ontology focuses on a different aspect of reality.(Chandrasekaran, Johnson, & Benjamins, 1999) Different domains do not need to express the fine details of other domains that they will never use. Even within a domain there might be multiple ontologies describing the same domain. For example Grosjean, Soualmia, Bouarech, Jonquet, and Darmoni (2014) mentions multiple portals

to search multiple ontologies in the Biomedical domain.

It might be of interest that we can distinguish between *Internal Interoperability* and *External Interoperability*. *Internal Interoperability* relates to interoperability between elements of a system. *External Interoperability* refers to the interoperability between systems. (García-Castro & Gómez-Pérez, 2011) We will be focussing on *External Interoperability*

If it is not possible for systems to use the same ontology, some interoperability can be achieved by using the same "upper-level" ontology. "A top-level ontology describes very general concepts that are the same for every domain and are not dependent on task and purpose." (van Diggelen & Dignum, 2007)

Another option is *Ontological Alignment*. This involves creating mappings between the entities and relationships of the ontologies. The problem with this approach is the sheer number of mappings that need to be made. (van Diggelen & Dignum, 2007) An intermediate ontology can be used to reduce the number of mappings. (Ciocoiu, Gruninger, & Nau, 2000; van Diggelen & Dignum, 2007)

When some form of *semantic interoperability* is achieved between systems, these systems can be used to answer queries. This can be done through *Ontology based data access* (OBDA). "Ontology-based data access (OBDA) is a recent paradigm that proposes the use of an ontology as a conceptual, reconciled view of the information stored in a set of existing data sources." (Bienvenu & Rosati, 2015) In other words the queries are done through the ontology to the data. This data could be anything from a relational database, triple stores or datalog engines. (Rodríguez-Muro, Kontchakov, & Zakharyashev, 2013)

These queries can be formally expressed using Descriptive Logics which is a family of formal knowledge representation languages. There are different types queries. Some types include *Conjunctive Queries* and *Instance Queries*. Bienvenu and Rosati (2015)

These queries in Descriptive Logics format then needs to be translated into the language of the target system. For instance if the target systems use a relational database, the query needs to be translated to a SQL query. (Kontchakov, Rodríguez-Muro, & Zakharyashev, 2013)

The Semantic Web is an extension of the the current Web, Berners-Lee et al. (2001), and allows for much more comprehensive searches than than traditional text based searches (Google,Bing). Allowing the user to specify not only the objects that they are searching for but the relationships these objects have.

## References

- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Sci. Am.*, 284(5), 34–43.
- Bienvenu, M., & Rosati, R. (2015). Query-based comparison of OBDA specifications. In *Ceur workshop proceedings* (Vol. 1350).
- Chandrasekaran, B., Johnson, T. R., & Benjamins, V. R. (1999). Ontologies: what are they? why do we need them? *IEEE Intelligent Systems and Their Applications*, 14, 20–26.
- Ciocioiu, M., Gruninger, M., & Nau, D. S. (2000). Ontologies for Integrating Engineering Applications. *Journal of Computing and Information Science in Engineering*, 1(1), 12–22.
- Decker, S., Melnik, S., Harmelen, F. V., Fensel, D., Klein, M., Broekstra, J., ... Horrocks, I. (2000). THE SEMANTIC WEB: The Roles of XML and RDF. *IEEE Internet Comput.*, 4(5), 63–74.
- García-Castro, R., & Gómez-Pérez, A. (2011). Perspectives in Semantic Interoperability. In *The semanticweb research and applications 8th extended semanticweb conference* (p. 16).
- Grosjean, J., Soualmia, L. F., Bouarech, K., Jonquet, C., & Darmoni, S. J. (2014). An Approach to Compare Bio-Ontologies Portals. In *Studies in health technology and informatics* (Vol. 205, pp. 1008–1012).
- Heiler, S. (1995). Semantic Interoperability. *ACM Computing Surveys*, 27(2), 271–273.
- Horrocks, I., Patel-Schneider, P., & van Harmelen, F. (2003). From SHIQ and RDF to OWL: The Making of a Web Ontology Language. *J. Web Semant.*, 1(1), 7–26.
- Kontchakov, R., Rodríguez-Muro, M., & Zakharyashev, M. (2013). Ontology-based data access with databases: A short course. In (Vol. 8067 LNAI, pp. 194–229).
- Rodríguez-Muro, M., Kontchakov, R., & Zakharyashev, M. (2013). Ontology-based data access: Ontop of databases. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 8218 LNCS(PART 1), 558–573.
- van Diggelen, J., & Dignum, F. (2007). Developing semantically interoperable e-commerce systems. In *Proceedings of the ninth international conference on electronic commerce* (pp. 117–126). New York, NY, USA: ACM.