AN ONTOLOGY TO SUPPORT NON-INVASIVE DIAGNOSIS OF HERITAGE METALS

Completed Research

De Santo, Alessio, HES-SO // University of Applied Sciences Western Switzerland, HEG Arc, Neuchâtel, Switzerland, alessio.desanto@he-arc.ch

Vonlanthen, Yann, HES-SO // University of Applied Sciences Western Switzerland, HEG Arc, Neuchâtel, Switzerland, yann.vonlanthen@he-arc.ch

Rosselet, Antoine, HES-SO // University of Applied Sciences Western Switzerland, HEG Arc, Neuchâtel, Switzerland, antoine.rosselet@he-arc.ch

Degrigny, Christian, HES-SO // University of Applied Sciences Western Switzerland, HE-Arc CR, Neuchâtel, Switzerland, christian.degrigny@he-arc.ch

Gaspoz, Cédric, HES-SO // University of Applied Sciences Western Switzerland, HEG Arc, Neuchâtel, Switzerland, cedric.gaspoz@he-arc.ch

Abstract

Metal conservators confronted with unknown artefacts rely on previous literature to develop treatment protocols. This search can be tedious given the dissemination of information across corpus of unstructured texts, mainly in the form of research papers and semi-structured databases of artefacts. In order to improve the search of artefacts sharing similar characteristics (metal composition and structure, conservation condition, etc.), this project proposes a hybrid search engine based on a domain ontology. Using a database populated with information resulting from comprehensive investigations of historic and archaeological artefacts, we extracted and selected key concepts and their relations through the use a various lexical analysis tools. Based on this corpus and frequency analysis, we were able to build an ontology of the domain, opening new perspective on information retrieval. Conservators are able to leverage the power of the hybrid search engine to compare their observations on a specific artefact with objects already stored in the database or with indexed research papers. Using keywords to describe corrosion forms they are confronted with, conservators can retrieve artefacts showing similar corrosion phenomena and assess the conservation condition of their artefacts, e.g. diagnosing the stability of metals or determining the location of the limit of the original surface in corrosion product crusts.

Keywords: stratigraphy, diagnosis, ontology, semantic, SPARQL.

1 The background of the problem

Heritage metal artefacts are found in various atmospheres (buried in the ground, submerged in the sea or exposed to indoor or outdoor pollutants). Their composition comprises antique metal elements such as Ag, Cu, Fe, Sn and Pb and modern ones such as Al, Mg, Ni and Zn. When these elements are combined with H, O, S and P, all sorts of compounds can be encountered. Understanding the forms of corrosion that develop on metal artefacts, and more particularly their active character, is a crucial task in the conservation process. This understanding can help in the choice of appropriate conservation protocols for working back to the limit of the original surface of the object while avoiding further deterioration by stopping the active corrosion processes. Working on heritage artefacts is a delicate task as they all have historic value. Therefore great importance is attached to ensure that the conservation techniques used will not damage them.

Bertholon (2000) has developed a methodology that provides conservators with a model of a stratigraphic representation of the corrosion layers. In that way, the various components constituting the corrosion form, from the core metal to the external corrosion layers, are depicted in strata (Figure 1). Each of the strata is of a specific nature and has multiple visual and non-visual characteristics. This methodology is used by conservators to locate the limit of the original surface of the object. It is expected then that appropriate conservation treatment is employed on the analysed artefact. Within the MiCorr application (Rosselet, Rochat, & Gaspoz, 2015), stratigraphies of the artefacts modelled with Bertholon's stratigraphic representation methodology can easily be compared, as they all follow the same modelling rules. Consequently, a conservator who works on an unknown metal artefact in a specific conservation condition can search for similar stratigraphies in a database of analysed artefacts. Such database currently exists in the forms of a report and a website. Moreover, there is also a need to retrieve comparable heritage artefacts based on keyword search, as some artefacts are not yet modelled with Bertholon's methodology. Also, metadata such as the origin and the environment in which the artefact was buried is not depicted in the stratigraphic representation and needs to be accessible through keyword search.

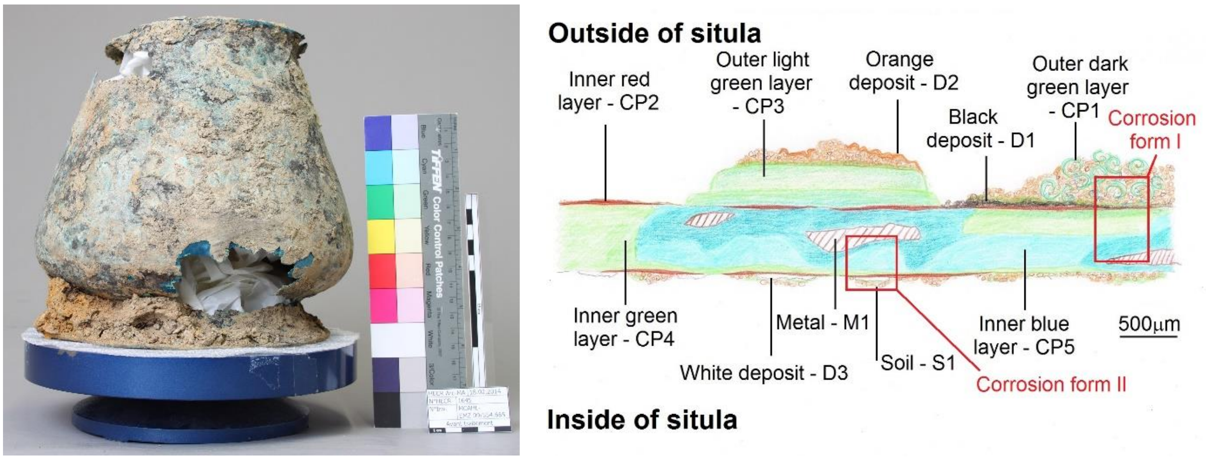

Figure 1: On the left a Celtic situla from the La Tène D period (140–30 BC) excavated from the Mormont sanctuary, La Sarraz/Eclépens, Vaud, Switzerland (Dudan 2009), Musée cantonal d'archéologie et d'histoire, Lausanne, © HE-Arc CR. On the right, a schematic representation of a cross-section of the situla drawn after microscopic observation, © HE-Arc CR.

2 The problem itself

Conservators already have a tool at their disposal that allows them to represent an artefact's corrosion layers (Rosselet, Grosjean, Degrigny & Gaspoz, 2016). Moreover, this tool facilitates the comparison of the conservation condition of their objects with the ones that are in the database. Thus, once a stratigraphy is drawn, it can be compared with others, which can give clues to conservators on the corrosion that might develop within their own artefact. However, such a comparison is not pertinent for conservators who are at the first stage of their artefact characterization process; they may need to make a quick selection of other objects that share similarities with theirs without having to draw an entire stratigraphy. In parallel, conservators may also want to refine the results returned by a stratigraphic search with information and/or metadata that is not included in the stratigraphy drawing. Thus, the keyword search can address this problem by returning a list of similar artefacts based on the conservator's criteria. When searching for an artefact that shares similarities with the one being analysed, a conservator needs to be provided with the most accurate results possible. Indeed, it is of little use to get a list of heritage artefacts that are not or only slightly related to the ones being assessed. For now, a lot of structured (e.g. database tables) and unstructured data (e.g. long texts) is included on the website, which contains a database of patrimonial objects. So far, we have been able to easily retrieve relevant information when it comes to structured data. Indeed, it is straightforward to return artefacts that share a common attribute with each other. However, the nature of unstructured data makes it more difficult to query against. Thus, a method that could extract and make links between words would be of great interest for our application. Additionally, knowing the context of a word permits a comparison with words of the same family and therefore to return more accurate results.

Moreover, given the large corpus of research on heritage artefacts, only a fraction of them are represented and described in the database. Thus, in order to extend the pertinence of the database without the need to populate it with all available research results, we should be able to include results from scientific papers within our results. Giving access to pertinent scientific articles when searching for a specific keyword would be of considerable use for conservation professionals. Again, there is a need for a tool that can understand the context of a searched word in order to return the most accurate results possible and help the conservators in their queries.

3 A plan for solving the problem

Research on ontology is widespread in the information systems community, and its importance is being recognized in a multiplicity of research fields and applications areas, including knowledge engineering, database design and integration and information retrieval and extraction (Guarino, 1998). The term 'ontology' tends to remain a bit vague, as it is used in very different ways (Guarino & Giaretta, 1995). In computer sciences, ontologies draw their origin in the Semantic Web (Berners-Lee & Fischetti, 2000). The World Wide Web Consortium (W3C) founded by the same Berners-Lee describes the Semantic Web as a 'web of data' understandable by machines, compared with the current 'web of documents' that machines simply display. An ontology is designed not only to provide a complete view of domain concepts but also to identify quickly and accurately similarities between concepts (Gómez-Pérez, Fernández-López & Corcho, 2004), even if not identical, and to conduct consistent alignments (Bedini & Nguyen, 2007). The ultimate goal is to enable computers to do more useful work and to develop systems that can support trusted interactions over the network. A simple example is the research results provided by hybrid semantic search engines (Hai Dong, Hussain & Chang, 2008) that offers a direct answer to the query without the need of visit a collection of returned links. An ontology is not only a classification, or taxonomy of general concepts, but is also a model that includes and maintains the most common properties of concepts, their relationships' existing alignments and known semantics (Bedini & Nguyen, 2007). An ontology would allow the storage of unstructured knowledge about the artefacts by highlighting concepts and their relations. Every artefact being unstructured and unique there is no structure that could store every specificity. To overcome this peculiarity, semantic structuration of the information through an ontology could be established. The creation of an ontology is a very difficult and time-consuming task (Drumond & Girardi, 2008). Research is putting efforts into fully, or at least partially, automating the ontology generation process. This field is best known as ontology learning. We can differentiate four categories of techniques (Table 1) (Bedini & Nguyen, 2007).

Conversion or translation	Starting from structured or semi-structured data and converting it into readable data for ontologies
Mining based	Starting from unstructured data and using text-mining techniques to extract knowledge
External knowledge based	Starting from external knowledge resources and extracting necessary knowledge
Frameworks	Using several techniques and tools to generate an ontology

Table 1. Ontology generation classification

When faced with a semi-structured database including unstructured data, there are three ways of building an ontology that can be distinguished. The first method would be to manually process all the data looking for concepts. This first method is the most effective, but also the most time-consuming. To ensure the quality of the ontology, manual processing should only be conducted by an expert in the particular field. The second method is to use software to automate the extraction of concepts by processing with a text mining algorithm. This solution is a lot faster with a large amount of data, but the quality of the ontology is not assured. The third solution combines the advantages of both the previous solutions: process the data with a text mining algorithm to create a first draft of the ontology and then have it validated and tested by an expert in the field.

Text mining – also known as text data mining or knowledge discovery from textual databases – refers to the process of extracting interesting and non-trivial patterns or knowledge from text documents (Tan, 1999). Two different approaches are employed – statistical and linguistic. While statistical approaches often rely on word frequencies and word co-occurrences, linguistic approaches make use of natural language processing techniques, such as syntactic, morpho-syntactic, lexico-syntactic and syntactic-semantic analysis, for extracting information (Drumond & Girardi, 2008). The best way to get relevant results is to find the right combination of these two approaches.

Most methods automate only some steps of the ontology generation process. To generate an ontology, there is still a lot of work that can scarcely be automated. In most cases, an ontology is not a static behaviour of a domain; we should be able to guarantee the natural evolution of it. Once an ontology is generated, we should be able to infer some logical consequences from a set of explicitly asserted facts or axioms. A reasoner can help us in this task and typically provides automated support for reasoning tasks such as classification, debugging and querying (Abburu, 2012). A reasoner will also check the consistency of the ontology.

4 The application of the solution

The MiCorr project contains two main sources of data: a database describing artefacts and a glossary of terms and definitions of the main concepts used in the field of conservation-restoration. In the current state of the database, each artefact is represented as a record with several sections following the structure shown in Table 2.

Structure
The object
Description and visual observation
Zones of the artefact submitted to visual observation and location of sampling areas
Macroscopic observation
Sample
Analyses and results
Metal
Corrosion layers
Synthesis of the macroscopic/microscopic observation of corrosion layers
Conclusion
References

Table 2. Artefact record structure

Some of the sections contain the artefact's attributes, which can be shared across multiple entities – through foreign key references – whereas others consist of plain text. Furthermore, images and tables add precision and refinement to the artefact description.

Searching for artefacts in the database can be performed with defined keywords which consist of the artefact's attributes. This already constitutes the starting point of an ontology, as the search tool knows that a specific attribute is linked to a characteristic of a defined nature. It is therefore able to return pertinent results based on the given keyword. However, the description of the artefacts also consists of plain text, which is difficult to query. That is why the information needs to be structured to return better results when searching for similar artefacts. In addition, links need to be added between the concepts so that the ontology becomes more relevant.

Following Fernández-López and collegues (1997), we used a multi-step process consisting of specification, conceptualization, formalization, integration and implementation, along with maintenance, knowledge acquisition, documentation and evaluation, in order to create the ontology. However, the use of evaluation leads to an iterative process because evaluation can lead to new specifications and formalization.

The first version of our ontology was created using a tool called D2RQ, which allows the creation of custom dumps of the database in RDF format for loading into a resource description framework (RDF) store (Bizer & Cyganiak, 2012). The RDF store (Klyne, Carroll & McBride, 2014) is very useful as it can then be imported into an ontology editor such as Protégé (Gonçalves et al., 2016). The idea is to convert the tables of the database into classes with their instances so that they can be exploited in an ontology. This is an automatic ontology learning technique described in (Michel, Montagnat & Faron-Zucker, 2014) which follows the form of a conversion/translation.

During the second step of the ontology creation, Protégé was used for manual refinements and improvements. This first version of our ontology was indeed modified with Protégé to become more pertinent and usable. Thus, classes have been renamed, and links between attributes have been added, to reflect the knowledge of the conservation-restoration field more precisely. Additionally, we added rules that prevent an instance from belonging to multiple classes and relations between instances. In order to improve this ontology, we collaborated closely with conservation professionals, who helped us to define the links between concepts and words in their field. At this stage, we performed several iterations to achieve a high level of satisfaction, providing an ontology that could be exploited. At the end of the project, the generated ontology included the following statistics (Table 3).

Classes	12
Individuals	186
Attributes	5
Axioms	373

Table 3. Generated ontology statistics

5 Utilisation

An ontology can be used in different ways to improve the diagnosis of heritage metals. The first use of the ontology is to improve the traditional full-text search engine. Hybrid semantic search engines combine traditional keyword-based search engines with semantic web technology (Hai Dong et al. 2008). In our case, we integrated the generated ontology to an existing search tool. We improved the effectiveness of a search box by providing some additional ontological functionalities. The search box is unmodified in appearance, but implements additional intelligence. The ontology is used to assist the research process, helping users in their task by providing them with better results that could assist them in diagnosing their artefacts. Therefore, users are supplied with advanced functionalities without changing their habits and landmarks. The most visible utilization of the ontology in our advanced search tool is our ability to offer autocompleting while entering a search term. The ontology proposes related words when a first word is typed into the search box (Figure 2).

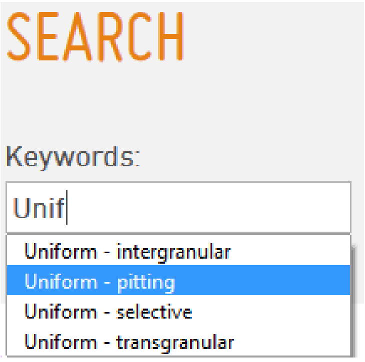

Figure 2. Autocompleting research example

On the second iteration, after displaying a first set of results, the user is again supplied with some related terms in order to refine the research. Object properties that link individuals of the same class permit the retrieval of connected elements. For instance, when the word Knife (an individual of TypeOfObject) has been researched and a set of corresponding artefacts is showed, related terms like Weapon and Household implement are suggested. Adding them to the research criteria will refine the set of results. The selection of linked individuals is the result of a SPARQL request. Other links established in the generated ontology allows the display of several other related terms in the same way as described above. However, the user is also supplied with related technologies, periods or corrosion forms and types.

In addition to returning pertinent artefacts from the database, the ontology can also be used to retrieve articles related to the keyword search. After parsing a collection of research papers on the topic of conservation-restoration, we are able to return references from articles about artefacts from the database when users are searching for specific terms. This is highly valuable for conservation professionals, because they can get more information from a single search, even if the artefact is not described in the database. Therefore, if the original search does not return pertinent results from the database, articles that contain the searched keyword will be returned to help conservators with their queries. This is possible thanks to the implementation of the ontology, which adds context to search keywords.

Creating an ontology also opens the field of plain text manipulation. Long texts contain words that can be found in the ontology and therefore a context can be added to them so that they can lead to more accurate results when searching. Furthermore, the ontology allows the gap between the two search tools on the website to be bridged. Once the ontology gets bigger, the characteristics of the strata added with the stratigraphic tool will be linked with the information in the textual database, allowing for more precise content to be returned. Eventually, the returned results will help conservators to decide which conservation methods to use for their artefacts.

6 The results

There are many techniques that can be used in order to evaluate the pertinence and validity of an ontology. Surveys show that most approaches to evaluation fall into one of four categories (Brank, Grobelnik & Mladenic, 2005): comparison with an existing ontology, evaluation of the results in an application, comparison with a corpus of texts from the same domain, or human expert evaluation. In the context of this research, we can only test the results of the ontology in an application, as well as test the ontology against a corpus of texts from the same domain. Indeed, given the absence of other ontologies from the same domain, there are no gold standards to compare our ontology with. We also discarded an evaluation based on expert feedback because we had already worked with some experts during the process of creating the ontology, which would undeniably introduce bias into the evaluation.

The first evaluation, using a corpus of texts from the domain, was performed in order to assess the completeness of the ontology regarding the concepts and instances that are included in it, as well as the correctness of the relations between the concepts. A sample of 13 scientific articles presenting researches on various artefacts was drawn from a research database and used as a text corpus to evaluate the ontology. The aim of the experiment was to establish some statistics in order to determine the completeness of the generated ontology. The experiment was run using the textual analysis tool of KNIME (Berthold et al., 2008). A collection of words was generated from the ontology and from the articles. We then computed the degree of matching between the two corpuses. This provided us with an assessment of on the coverage of the ontology on the recurrent concepts of the conservation-restoration field.

In order to generate the collection of words, we parsed the OWL / RDF file of the ontology with an XML reader in order to obtain the names of the individuals. The number of words listed was 275. We then used a PDF parser to list the words used in the articles (Table 4), and the redundancies were eliminated in all the lists.

Articles	**Number of matching words**
Alvarez, 2013	64
Bouchar, Dillmann & Neff, 2013	37
Cameron, Greaves, Northover & Connor, 2013	47
Cano, Iglesia, Lafuente, Bastidas & Navarro, 2013	31
Carlson, Lipfert, Ronnberg & Scott, 2010	49
Chiavari et al., 2013	29
Emmerson & Watkinson, 2013	35
Gillies & Seyb, 2013	42
Koleini, Prinsloo, Schoeman, Pikirayi & Chirikure, 2013	46
Marchand et al., 2013	31
Northover, Northover & Wilson, 2013	40
Scott & Maish, 2010	74
Wang, Huang & Shearman, 2009	57

Table 4. Statistics per article

There were 148 words of the ontology that appeared at least once in the articles, representing more than 53% of the ontology content. On average, 45 words of the ontology, representing more than 16% of the ontology content, were found in all articles. Moreover, 71% of the words with the highest frequency among the text corpuses were present in the ontology. Considering that the articles often present a small fraction of the domain knowledge, we felt that the current coverage of the ontology was satisfying.

The second test performed in order to evaluate the ontology is an application of the ontology for retrieving artefacts created using the same technology. The technology used to shape the metal depends on the period and the type of metal alloy, and this can be inferred from the microstructure of the metal. Therefore, after identifying the type of microstructure, the user is able to use the ontology to infer the possible technologies used to create the artefact. From this list of possible technologies, we can retrieve artefacts made with the same techniques.

Using SPARQL requests we could, for example, find all techniques used to shape metallic artefacts presenting a ‘dendritic structure with inclusions’. This can be used in a standalone request or be used to refine existing requests. The results returned can then be evaluated in order to assess the accuracy of the ontology in inferring properties of the artefacts. Despite the small number of artefacts in the database, we had very good results, but they are currently not significant owing to the fact that we were not able to use a different set of artefacts for the tests from the ones we used to create the ontology. Therefore, pending the addition of more artefacts to the database, we are not able to statistically validate this part of the test of the ontology.

```
SELECT ?searchedMicroStrURI ?technoURI ?technoLabel
WHERE {
?searchedMicroStrURI ont:resultsFromTechnology ?technoURI .
?searchedMicroStrURI rdfs:label ?searchedMicroStrLab .
?technoURI rdfs:label ?technoLabel .
FILTER(CONTAINS(UCASE(?searchedMicroStrLab), UCASE("Dendritic structure with
inclusions")))
}
ORDER BY ?technoLabel
```

Figure 3. Example request for inferring technologies from microstructures

In conclusion, these results encouraged us by demonstrating that the generated ontology properly covers the domain and that it can be successfully used to improve the quality of the comparison of artefacts based on their characteristics. Given the attractiveness of the database and its support for the diagnosis of ancient metallic artefacts, we expect to be able to further refine and improve the ontology with an increased number of additions to the database. On basis of these results, we can deduce that the generated ontology makes sense outside of its single use within MiCorr, and that it actually covers the main concepts of the diagnosis of ancient metallic artefacts within the conservation-restoration field.

Finally, all the work around the ontology generation was done in collaboration with an expert in the conservation-restoration field. No items were added as a result of analyses performed solely by software tools, this means that we can guarantee that the ontology contains only audited statements.

7 Conclusion

This research addresses an issue that lies at the intersection of two disciplines: conservation-restoration and information systems. The conservation-restoration researchers are confronted with an increased need of computer-aided systems in order to process and retrieve information from large unstructured corpuses of documents. From the other direction, information systems researchers are faced with the challenge of working with document corpuses from a field mostly foreign to them, but with the goal of identifying and extracting the most relevant information from them. This is the interdisciplinary component of this research which generated the most interesting challenges, but also the most rewarding ones.

Starting from various corpuses of information, we were able to extract and process their vocabulary in order to identify the main underlying concepts, attributes and relations of the conservation-restoration field of study. After multiple refinements with domain experts and statistical analysis, the generated ontology was tested for completeness and for its ability to make relevant inferences. Both evaluations brought positive conclusions, opening the way to the implementation of the ontology as a diagnosis tool. Although there is an increased number of applications offering semantic search engines based on such ontologies (Sudeepthi, Anuradha & Babu, 2012), we chose to follow a hybrid approach of combining both semantic and keyword-based search engines in order to hide the complexity of the tool for conservators. The resulting implementation improves the overall quality of the tool in offering better results to researchers looking for artefacts that present similar characteristics to the one they are studying. Given the restrictions on the analysis that can be performed on ancient artefacts, an improved non-invasive diagnosis tool is of great value for the conservation-restoration field and will ultimately lead to better conservation treatments. Indeed, if conservators are able to find and analyse treatments and their results from past restorations of metallic artefacts, they will be able to take more informed decisions about the treatment to apply to their artefact.

8 Acknowledgements

This paper is partly based on work supported by the RCSO ISnet under Award No. 38996. We would like to thank our colleagues, Valentin Boissonnas, Régis Bertholon and Romain Jeanneret from the conservation-restoration department of the Haute École Arc, who supported our work on this project with their continual feedback. We would also like to thank the students of the MA in conservation-restoration for participating in field tests of the tool.

References

Abburu, S. (2012). A Survey on Ontology Reasoners and Comparison. *International Journal of Computer Applications*, *57*(17), 33–39. http://doi.org/10.5120/9208-3748

Bedini, I. & Nguyen, B. (2007). Automatic Ontology Generation : State of the Art. *Evaluation*, 1–15.

Berners-Lee, T. & Fischetti, M. (2000). *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*.

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., Wiswedel, B. (2008). KNIME: The Konstanz Information Miner. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data Analysis, Machine Learning and Applications* (pp. 319–326). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-540-78246-9_38

Bertholon, R. (2000, December 20). *La limite de la surface d'origine des objets métalliques archéologiques. Caractérisation, localisation et approche des mécanismes de conservation.* Université Panthéon-Sorbonne - Paris I.

Bizer, C. & Cyganiak, R. (2012). D2RQ Platform – Accessing Relational Databases as Virtual RDF Graphs. Retrieved from http://d2rq.org/

Brank, J., Grobelnik, M. & Mladenic, D. (2005). A survey of ontology evaluation techniques. In M. Grobelnik & D. Mladenic (Eds.), *Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005)* (p. 4). Ljubljana, Slovenia.

Drumond, L. & Girardi, R. (2008). A survey of ontology learning procedures. *CEUR Workshop Proceedings*, *427*.

Fernández-López, M., Gómez-Pérez, A. & Juristo, N. (1997). METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In *AAAI-97 Spring Symposium Series* (Vol. SS-97–06, pp. 33–40). http://doi.org/10.1109/AXMEDIS.2007.19

Gómez-Pérez, A., Fernández-López, M. & Corcho, O. (2004). *Ontological engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web* (1st ed.). London, UK: Springer. http://doi.org/10.1007/b97353

Gonçalves, R., Hardi, J., Horridge, M., Musen, M., Nyulas, C., Tu, S. & Tudorache, T. (2016). Protégé. Retrieved from http://protege.stanford.edu/

Guarino, N. (1998). Formal Ontology and Information Systems. In N. Guarino (Ed.), *Formal Ontology in Information Systems* (pp. 3–15). Trento, Italy: IOS Press.

Guarino, N. & Giaretta, P. (1995). Ontologies and Knowledge Bases: Towards a Terminological Clarification. *Towards Very Large Knowledge Bases Knowledge Building and Knowledge Sharing*, *1*(9), 25–32. http://doi.org/10.1006/ijhc.1995.1066

Hai Dong, Hussain, F. K. & Chang, E. (2008). A survey in semantic search technologies. In *2008 2nd IEEE International Conference on Digital Ecosystems and Technologies* (pp. 403–408). Phitsanulok: IEEE. http://doi.org/10.1109/DEST.2008.4635202

Klyne, G., Carroll, J. J. & McBride, B. (2014). RDF 1.1 Concepts and Abstract Syntax. Retrieved from https://www.w3.org/TR/rdf11-concepts/

Michel, F., Montagnat, J. & Faron-Zucker, C. (2014). *A survey of RDB to RDF translation approaches and tools*. Nice: I3S.

Rosselet, A., Grosjean, M., Degrigny, C. & Gaspoz, C. (2016). Computer-Aided Support System for Metal Diagnosis of Patrimonial Objects. In *New Advances in Information Systems and Technologies* (pp. 961–971). http://doi.org/10.1007/978-3-319-31232-3_91

Rosselet, A., Rochat, V. & Gaspoz, C. (2015). Design of a new data structure to support non-invasive diagnostic on heritage metals. In *Proceedings of the 9th Mediterranean Conference on Information*

Systems (MCIS). Samos, Greece.

Sudeepthi, G., Anuradha, G. & Babu, M. (2012). A survey on semantic web search engine. *International Journal of Computer Science, 9*(2), 241–245.

Tan, A.-H. (1999). Text Mining: The state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Disocovery from Advanced Databases*, *8*, 65–70. http://doi.org/10.1.1.132.6973

Annexes

References used for the evaluation of the ontology.

Alvarez, R. (2013). Restoration and archeometallurgical study of a roman cauldron made of copper alloy. *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, *46*(Us 2070), 321–328.

Bouchar, M., Dillmann, P., & Neff, D. (2013). Atmospheric corrosion of iron-based reinforcement of gothic cathedrals - overview of possible corrosion systems and influence of the structure and composition of the corrosion layers on the corrosion diagnosis. *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, 215–221.

Cameron, E., Greaves, P., Northover, P., & Connor, S. O. (2013). Royal forteviot : the recovery , conservation strategy and analysis of a bronze age dagger from a high status cist burial. *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, 21–28.

Cano, E., Iglesia, A. La, Lafuente, D., Bastidas, D. M., & Navarro, J. V. (2013). Corrosion by sulphur of bronze reinforcing elements in León cathedral (Spain). *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, 209–214.

Carlson, M. O., Lipfert, N. R., Ronnberg, E. A. R., & Scott, D. A. (2010). Technical analysis of Muntz metal sheathing from the American clipper ship Snow Squall (1851-1864). *Metal 2010: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group, October 11-15, 2010, Charleston, South Carolina, USA*, *1862*, 107–115.

Chiavari, C., Bernardi, E., Cauzzi, D., Volta, S., Bignozzi, M. C., Lenza, B., … Martini, C. (2013). Influence of natural patinas of outdoor quaternary bronzes on conservation treatments. *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, 159–168.

Emmerson, N., & Watkinson, D. (2013). Preparing Historic Wrought Iron for Protective Coatings : *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, 119–128.

Gillies, J. C., & Seyb, I. (2013). La Fée aux Fleurs: investigation and conservation of a 19th century outdoor cast iron sculpture. *METAL 2013 Edinburgh, Scotland 16th – 20th September 2013, Interim Meeting of the ICOM-CC Metal Working Group*, 147–152.

Koleini, F., Prinsloo, L. C., Schoeman, M. H. A., Pikirayi, I., & Chirikure, S. (2013). Characterisation of the corrosion layer on iron archaeological artefacts from K2 (825-1120 AD), an archaeological site in South Africa. *Studies in Conservation*, *58* (3), 274–282. doi:DOI 10.1179/2047058412Y.274 0000000044

Marchand, G., Chevallier, R., Guilminot, E., Rossetti, L., Lemoine, S., & Vieau, M. (2013). Study of the conservation treatments applied to archaeological. *METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group*, 241–246.

Northover, P., Northover, S., & Wilson, A. (2013). Microstructures of ancient and historic silver.

METAL 2013: Proceedings of the Interim Meeting of the ICOM-CC Metal Working Group, 253–260.

Wang, Q., Huang, H., & Shearman, F. (2009). Bronzes from the Sacred Animal Necropolis at Saqqara, Egypt: a study of the metals and corrosion. *British Museum Technical Research Bulletin, 3*, 73–82.