

Editorial: Special Issue on Semantic Web for Cultural Heritage

Antonis Bikakis^a, Eero Hyvönen^b, Stéphane Jean^c, Béatrice Markhoff^d and Alessandro Mosca^e

^a *Department of Information Studies, University College London, London, United Kingdom*

E-mail: a.bikakis@ucl.ac.uk

^b *Helsinki Centre for Digital Humanities (HELDIG) and Aalto University, Finland*

E-mail: eero.hyvonen@aalto.fi

^c *LIAS, ISAE-ENSMA and University of Poitiers, Poitiers, France*

E-mail: jean@ensma.fr

^d *LIFAT, University of Tours, Blois, France*

E-mail: beatrice.markhoff@univ-tours.fr

^e *KRDB Research Centre, Free University of Bozen-Bolzano, Bolzano, Italy*

E-mail: alessandro.mosca@unibz.it

Editors: Pascal Hitzler, Kansas State University, USA; Krzysztof Janowicz, University of California, Santa Barbara, USA

Abstract. Cultural Heritage and Digital Humanities have become major application fields of Linked Data and Semantic Web technologies. This editorial introduces the special issue of the Semantic Web (SWJ) journal on Semantic Web for Cultural Heritage. In total 30 submissions for the call of papers were received, of which 11 were selected for publication. The papers cover a wide spectrum of modelled topics related to language, reading and writing, narratives, historical events and cultural artefacts, while describing reusable methodologies and tools for cultural data management. This issue indicates and demonstrates the high potential of Semantic Web technologies for applications in the Cultural Heritage domain.

Keywords: Semantics, Ontologies and information modelling, Data access and exploration, Knowledge Graphs and Linked data, Cultural Heritage, Digital Humanities

1. Introduction

Cultural Heritage plays a central role to better understand previous generations and the history of where humankind comes from, and to envision where it is going to. The Web allows people to publish, explain, and debate at all scales, local, national, and worldwide. Scientific researchers, organisations, associations, and schools are looking for relevant technologies for annotating, integrating, sharing, accessing, visualising, and analysing the mine of cultural collections and, more generally, cultural data. There is also a need for taking into account profiles and preferences of end users in order to offer them highly personalised digital experiences. Several national and international research,

innovation, and infrastructure programmes, such as EUROPEANA [1], DARIAH [2], PARTHENOS [3], CrossCult [4], ARCHES [5], and ARIADNEplus [6], have been launched to these directions. During 2018, which was the European Year of Cultural Heritage [7], several events and initiatives across Europe encouraged people to engage, explore and debate our rich and diverse Cultural Heritage.

When dealing with scholarly data, the “FAIR guiding principles for scientific data management and stewardship” [8] of publishing Findable, Accessible, Interoperable, and Re-usable data are a common norm. A fundamental challenge that many of the aforementioned projects deal with is how to make Cultural Heritage data, which is made available by different actors

1 in different cultural domains and in a multitude of dif-
2 ferent languages and formats, mutually interoperable,
3 so that it can be searched, linked, and presented in a
4 harmonised way across the boundaries of the datasets
5 and data silos.

6 Early solutions were based on the syntactic or struc-
7 tural level of data, without leveraging the rich semantic
8 structures underlying the content. During the two last
9 decades, solutions based on the principles and tech-
10 nologies of the Semantic Web have been proposed to
11 explicitly represent the semantics of data sources and
12 make both their content and their semantics machine
13 operable and interoperable. In parallel, knowledge rep-
14 resentation models have matured, such as the CIDOC-
15 CRM's ecosystem of the museum sector, interlinked
16 with FRBR-based models in libraries, which is dedi-
17 cated to the cultural heritage area including the fields
18 of documentation, archaeology, history, architecture,
19 etc. As more and more institutions bring their data
20 to the Semantic Web level, the tasks of data integra-
21 tion, sharing, analysis, visualisation, etc. are now to
22 be conceived in this very rich framework. At the same
23 time, Artificial Intelligence based methods are increas-
24 ingly used both in semantic content creation and in
25 the development and support of applications for human
26 users.

27 This special issue has offered to Computer Sci-
28 entists, Data Scientists and Digital Humanities re-
29 searchers who are involved in the development or
30 deployment of Semantic Web solutions for Cultural
31 Heritage the opportunity to present their realisations,
32 the outcomes of their projects, being either publicly
33 reusable Semantic Web tools, datasets or ontologies
34 published in the Linked Open Data Cloud, or Sema-
35 ntic Web techniques, services and architectures for Cul-
36 tural Heritage.

37 2. Special issue contributions

38
39 From the large variety of the 30 submissions coming
40 from 16 different countries, the following papers were
41 accepted to be included in the special issue at hand.
42 The papers are classified in three categories, according
43 to the main area or domain of contribution.

44 2.1. Language, Reading, Writing

45
46
47 The paper “*Ce qui est écrit et ce qui est parlé. CRMtex for modelling textual entities on the Semantic Web*” presents CRMtex, an ontology for modelling

1 texts, with an emphasis on its more recent develop-
2 ment. It describes the design rationale of the ontol-
3 ogy, its classes and properties and their relations to el-
4 ements of CIDOC-CRM. As demonstrated with exam-
5 ples, the proposed ontology is able to model not only
6 texts as physical and linguistic entities, but also activi-
7 ties and procedures related to them, such as their pro-
8 duction, transcription and decoding. Built as an exten-
9 sion of CIDOC-CRM, CRMtex achieves its aim of be-
10 ing an interoperable data model, which can be used for
11 different types of texts and for different purposes.

12 The paper “*Modeling Execution Techniques of In-*
13 *scriptions*” complements the previous article by focus-
14 ing on the description of writing execution techniques.
15 The limitations of the EAGLE and CRMtex models for
16 this kind of information are explained and addressed
17 by the authors' proposal, which can be combined with
18 these models. Interestingly, this ontology (again ex-
19 tending CIDOC-CRM) complements EAGLE Vocabu-
20 laries (expressed in SKOS) with a class structure.

21 In “*Understanding the phenomenology of reading*
22 *through modelling*”, the authors address another hu-
23 man activity, strongly related to texts or inscriptions:
24 reading. It discusses the design process of an ontol-
25 ogy that models the human's experience of reading,
26 called READ-IT, carried out by an international multi-
27 disciplinary research team. This ontology is meant to
28 semantically annotate sources of studies about reading
29 events. As in most of the approaches presented in this
30 special issue, the CIDOC-CRM is extensively reused
31 in READ-IT. The resulting ontology is available in a
32 GitHub repository.

33 2.2. Narratives, History, Archaeology

34
35 The paper “*Of Lions and Yakshis: Ontology-based*
36 *Narrative Structure Modelling for Culturally Diverse*
37 *Folktales*” describes an ontology for folk tales, which
38 is based on Vladimir Propp's theory “Morphology of
39 the Folktale”. The aim of the ontology is to assist the
40 analysis of folk tales by humanities researchers. The
41 paper describes the data modelling approach, the de-
42 sign and the implementation of the ontology. It also
43 presents a tool for semi-automatic extraction of infor-
44 mation from folk tales (in textual form) and describes
45 how the proposed ontology was applied for the analy-
46 sis and comparison of African and Indian folk tales.

47 The notion of narratives is at the core of the pa-
48 per “*Representing Narratives in Digital Libraries: The*
49 *Narrative Ontology*”, which presents a general ap-
50 proach for modelling narratives. It introduces a formal
51

1 expression of the concept of “narrative” and the re-
2 sulting “Narrative Ontology” (NOnt) is the first-order
3 logic-based counterpart of this expression. NOnt is an
4 extension of well-known standards like CIDOC-CRM,
5 FRBRoo and the W3C Time Ontology, and it is cur-
6 rently implemented with the SWRL rule language. The
7 formalisation effort behind the development of NOnt
8 has also given rise to the implementation of a semi-
9 automatic software, named “Narrative Building and
10 Visualising Tool” (NBVT), to create narratives and vi-
11 sualise them in several ways. The implementation of
12 this tool is based on Semantic Web technologies and its
13 main features are also presented in this contribution.

14 The paper “*WarSampo knowledge graph: Finland in*
15 *the Second World War as Linked Open Data*” presents
16 a shared knowledge graph, semantic infrastructure,
17 and Linked Open Data service for publishing data
18 about the World War II. The knowledge graph and data
19 service have been used, e.g., for implementing the in-
20 use semantic portal “*WarSampo–Finnish WW2 on the*
21 *Semantic Web*” that has had hundreds of thousands of
22 users on the Web. The system is based on represent-
23 ing war as a spatio-temporal sequence of events that
24 soldiers, military units, and other actors participate in
25 different roles. To support sustainability of the knowl-
26 edge graph, a data transformation and linking pipeline
27 has been created. The WarSampo knowledge graph,
28 totalling approximately 14 million triples, is openly
29 available as a service on the Linked Data Finland plat-
30 form, and is part of the international LOD Cloud.

31 The paper “*A challenge for historical research:*
32 *making data FAIR using a collaborative ontology man-*
33 *agement environment (OntoME)*” argues that the appli-
34 cation of the FAIR data principles in the field of his-
35 torical research requires the development and use of a
36 standard ontology. It proposes adopting CIDOC-CRM
37 as the core ontology for this domain, in combination
38 with two other foundational ontologies, C.DnS and
39 DOLCE. It also argues about the need of a collabora-
40 tive web environment, which will enable researchers to
41 commonly develop specifications of the core ontology
42 for specific sub-domains or applications and align the
43 different metadata models used by different projects.
44 Finally it explains how the ontology management en-
45 vironment OntoMe can serve this purpose.

46 In the field of archaeology, the paper “*OntoAn-*
47 *dalus: an ontology of Islamic artefacts for terminologi-*
48 *cal purposes*” adopts the same foundational ontologies
49 (DOLCE+DnS Ultralite (DUL)) as the previous one.
50 It presents an ontology for Andalusian pottery arte-
51 facts, called OntoAndalus, built as a specialisation of

1 DOLCE+DnS Ultralite (DUL) and modelled in OWL.
2 Its development relied on the interpretation of a corpus
3 from Portuguese and Spanish domain specific texts,
4 English textbooks and reference works, as well as from
5 more specialised documents from related conferences
6 and journals articles. The paper describes the main de-
7 sign patterns regarding the modelling of artefact types,
8 events and tasks, and uses the case study of *Vaso de*
9 *Tavira* to exemplify how these patterns were applied to
10 model lighting artefacts, the life cycle of pottery and
11 the several descriptions of the artefact.

12 2.3. Tools for Data Management: Designing, 13 Querying, Analysing 14

15
16 The ideas and approach of the “*Pattern-based de-*
17 *sign applied to cultural heritage knowledge graphs*”
18 paper are rooted in the lessons learned, the method-
19 ologies and the modelling choices discovered during
20 the development of ArCo, a knowledge graph consist-
21 ing of a network of (EDM and CIDOC-CRM-aligned)
22 ontologies that model the Cultural Heritage domain
23 and a Linked Open dataset of around 172.5M triples
24 about Italian cultural properties. The paper argues
25 about the advantages of embracing the “eXtreme De-
26 sign (XD)” in the creation of Cultural Heritage ontolo-
27 gies, a methodology inspired by the Extreme Program-
28 ming (XP) software development approach. It provides
29 the details behind the modelling of the ArCo ontol-
30 ogy network, the architectural patterns in place and
31 the characteristics of the evaluation that has been per-
32 formed.

33 The paper “*Applying and Developing Semantic Web*
34 *Technologies for Exploiting a Corpus in History of Sci-*
35 *ence: the Case Study of the Henri Poincaré Correspond-*
36 *ence*” presents a semantic virtual research environ-
37 ment dedicated to Henri Poincaré’s letters digital cor-
38 pus, i.e. letters with descriptive, scientific and mathe-
39 matical content. Semantic Web technologies are used
40 to enhance both annotation and querying of this cor-
41 pus. Concerning the semantic annotation, RDFS en-
42 tailment is leveraged to propose a ranked list of po-
43 tential values for the RDF triples associated to specific
44 parts of the letters. For querying, transformation rules
45 on SPARQL queries are defined to support approxi-
46 mate searches on vague concepts such as “*the end of*
47 *the 19th century*”, which is a recurrent need in the Cul-
48 tural Heritage context.

49 Another correspondence of a mathematician is one
50 of the use cases of Gravsearch, a system that sup-
51 ports complex searches in virtual graphs, introduced

1 in “*Gravsearch: transforming SPARQL to query hu-*
 2 *manities data*”. Gravsearch is a SPARQL query rewrit-
 3 ing system which aims at supporting both the develop-
 4 ers and the users with the introduction of an abstrac-
 5 tion layer on top of the existing triplestore implemen-
 6 tations. Gravsearch has been developed as part of the
 7 Knora (Knowledge Organization, Representation, and
 8 Annotation) API, an application by the “Data and Ser-
 9 vice Center for the Humanities” (DaSCH) whose main
 10 focus is on the preservation and promotion of digital
 11 data in the Humanities through dedicated data manage-
 12 ment methods, data storage solutions and data access
 13 platforms.

14 3. Summary and future directions

15
 16
 17
 18 The variety and large number of papers submitted
 19 to this special issue suggest that Linked Data and Se-
 20 mantic Web technologies are becoming increasingly
 21 important in creating, publishing, and analysing Cul-
 22 tural Heritage data in Digital Humanities. As more and
 23 more data is becoming available in harmonised inter-
 24 operable datasets, more and more intelligent applica-
 25 tions for searching, exploring, and analysing semanti-
 26 cally structured cultural data are also emerging. The
 27 contributions in this special issue witness this develop-
 28 ment.

29 Another interesting observation is that all papers
 30 in this special issue, as well as the vast majority of
 31 the papers that were not accepted, rely on existing
 32 standard ontologies and data models for the repre-
 33 sentation and interchange of cultural data, such as
 34 CIDOC-CRM, FRBR and EDM. The usage of upper
 35 level ontologies (e.g., DOLCE), ontology design pat-
 36 terns (e.g., c.DnS), and domain-independent standards
 37 (e.g., SKOS and DCMi), is also well testified. In this
 38 way they adopt one of the re-usability principles of
 39 FAIR data, which recommends that “(meta)data meet
 40 domain-relevant community standards”. But also, very
 41 importantly, they ensure the soundness and quality of
 42 their approaches, and exploit the methods and tools
 43 that have been developed for such models.

44 Although several significant steps have already been
 45 made, there are still several hindrances for fulfilling
 46 the potential of Semantic Web technologies in Cultural
 47 Heritage. One of them is that, although there are now
 48 well-established ontologies for the Cultural Heritage
 49 domain and most of the related fields, there are only
 50 few tools that humanities scholars, museum practition-
 51 ers and other people working in this domain, can easily

1 use to model, manage, analyse and interlink cultural
 2 data. Some of the papers presented in this special issue
 3 attempt to address this problem and present some very
 4 promising results. There is still, though, much more
 5 work to be done, and an important lesson learned from
 6 the introduced projects, is that the involvement of the
 7 people who work in this domain both in the design and
 8 the evaluation of such tools is essential for ensuring
 9 that they will fulfill their design purposes.

10 As it is also obvious from the papers presented in
 11 this special issue, most of the current research in this
 12 field is still led by universities or research institutes,
 13 some of them in collaboration with large cultural her-
 14 itage organisations. Smaller organisations are left be-
 15 hind. According to two studies on the adoption of open
 16 data practices [9, 10] (which are closely related with
 17 the FAIR and Linked Data ones), the main challenges
 18 that they face are the extra time, efforts and costs re-
 19 quired for the digitisation of their collections, their
 20 proper documentation and rights clearance; the lack of
 21 metadata for their collections; and the lack of relevant
 22 skills among their staff. Of course, the digitisation of
 23 their collections is not a problem that Semantic Web
 24 technologies can solve. However, Semantic Web re-
 25 searchers can help alleviate some of these challenges
 26 by developing and providing well-documented tools
 27 and detailed guidelines on how the FAIR or Open or
 28 Linked Data principles can be applied, but also plat-
 29 forms that will enhance the communication, informa-
 30 tion exchange, collaboration and networking between
 31 the cultural institutions.

32 Acknowledgements

33
 34
 35
 36 *We would like to thank the Editors-in-Chief Pascal*
 37 *Hitzler and Krzysztof Janowicz for their support in*
 38 *compiling this special issue. The following list includes*
 39 *the names of the reviewers (in alphabetical order) who*
 40 *participated in the review process of the manuscripts*
 41 *submitted to the special issue. Their effort and excel-*
 42 *lent work have been fundamental for the publication of*
 43 *the high quality scientific contributions composing the*
 44 *present Semantic Web Journal volume.*

45 Reviewers: Trond Aalberg, Valentina Bartalesi,
 46 Sotiris Batsakis, Stefano Borgo, Carmen Brando,
 47 George Bruseker, Giannis Chantas, Brice Chardin,
 48 Benjamin Cogrel, Roberto Confalonieri, Olivier
 49 Corby, Gayo Diallo, Oyvind Eide, Achille Felicetti,
 50 Roberta Ferrario, Géraud Fokou, Pietro Galliani,
 51 Pawel Garbacz, Andrea Giovanni Nuzzolese, Günther

1 Görz, Anais Guillem, Allel Hadjali, Fayçal Hamdi,
2 Mikko Koho, Kalliopi Kontiza, Efstratios Kontopou-
3 los, Nikolaos Lagos, Pietro Liuzzo, Martin Lopez
4 Nores, Claudio Masolo, Francesco Massucci, Carlo
5 Meghini, Albert Meroño, Amedeo Napoli, Yannick
6 Naudet, Franco Niccolucci, Laura Pandolfo, Rafael
7 Peñaloza, Tiago Prince Sales, Cedric Pruski, Mar-
8 tin Rezk, Catherine Roussey, Guillem Rull, Emilio
9 M. Sanfilippo, Daniele F. Santamaria, Christoph
10 Schlieder, Michalis Sfakakis, Sofia Stamou, Ranka
11 Stankovic, Maria Theodoridou, Konstantin Todorov,
12 Jouni Tuominen, Yannis Tzitzikas, Genoveva Vargas-
13 Solar, Fabio Vitali, Andreas Vlachidis, Holly Wright.

14 References

- 15
16
17
18 [1] The EUROPEANA web portal. <https://www.europeana.eu/>.
19 [2] The Digital Research Infrastructure for the Arts and Humanities (DARIAH). <https://www.dariah.eu/>.

- [3] Pooling Activities, Resources and Tools for Heritage
E-research Networking, Optimization and Synergies
(PARTHENOS). <https://www.parthenos-project.eu/>.
[4] The CrossCult web portal. <https://www.crosscult.eu/>.
[5] The web portal of the ARCHES Platform. <https://www.archesproject.org/>.
[6] The ARIADNEplus web portal. <https://ariadne-infrastructure.eu/>.
[7] The European Year of Cultural Heritage 2018. https://europa.eu/cultural-heritage/european-year-cultural-heritage_en.html.
[8] The FAIR Guiding Principles for scientific data management and stewardship. <https://www.go-fair.org/fair-principles/>.
[9] E. Foundation, Transforming the world with culture: Next steps on increasing the use of digital cultural heritage in research, education, tourism and the creative industries, Technical Report, Europeana Foundation, 2015.
[10] B. Estermann, Diffusion of Open Data and Crowdsourcing among Heritage Institutions: Based on data from Finland, Poland, Switzerland, and The Netherlands, in: *EGPA 2015, held on 26-28 August 2015 in Toulouse, France, 2015*.