Technical usability of Wikidata's linked data

Evaluation of machine interoperability and data interpretability

Nuno Freire[1[0000-0002-3632-8046]] and Antoine Isaac[2,3[0000-0001-9767-6979]]

[1] INESC-ID, Lisbon, Portugal
[2] Europeana Foundation, The Hague, The Netherlands
[3] Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
nuno.freire@tecnico.ulisboa.pt,
antoine.isaac@europeana.eu

Abstract. Wikidata is an outstanding data source with potential application in many scenarios. Wikidata provides its data openly in RDF. Our study aims to evaluate the usability of Wikidata as a data source for robots operating on the web of data, according to specifications and practices of linked data, the Semantic Web and ontology reasoning. We evaluated from the perspective of two use cases of data crawling robots, which are guided by our general motivation to acquire richer data for Europeana, a data aggregator from the Cultural Heritage domain. The first use case regards general data consumption applications based on RDF, RDF-Schema, OWL, SKOS and linked data. The second case regards applications that explore semantics relying on Schema.org and SKOS. We conclude that a human operator must assist linked data applications to interpret Wikidata's RDF because of the choices that were taken at Wikidata in the definition of its expression in RDF. The semantics of the RDF output from Wikidata is "locked-in" by the usage of Wikidata's own ontology, resulting in the need for human intervention. Wikidata is only a few steps away from high quality machine interpretation, however. It contains extensive alignment data to RDF, RDFS, OWL, SKOS and Schema.org, but a machine interpretation of those alignments can only be done if some essential Wikidata alignment properties are known.

Keywords: RDF, RDFS, OWL, Schema.org, Semantic Web.

1 Introduction

Wikidata is an outstanding data source with potential application in many scenarios. Wikidata provides its data openly in RDF. Our study aims to perform an evaluation of the usability of Wikidata as a data source for robots operating on the web of data, according to specifications and practices of linked data and the semantic web. We evaluated from the perspective of two use cases of data crawling robots. The first use case regards general data consumption applications (potentially with Big Data requirements) based on RDF, RDF-Schema, OWL and linked data. The second case regards cross-domain applications that explore semantics relying on general-purpose, shared vocabularies like SKOS and Schema.org.

These uses cases are motivated by our long-term goal of using linked data as a source for descriptions of cultural heritage resources by large aggregators of cultural heritage data, in particular for Europeana[1]. Europeana has the role of facilitating the usage of digitized cultural heritage resources from and about Europe [1]. Although many European cultural heritage institutions do not yet have a presence in Europeana, it already holds metadata from over 3,700 providers, mostly libraries, museums and archives. Some of our early exploratory research on Wikidata has shown that it can be a rich source for data on cultural heritage and for digital representations of the cultural heritage objects (images, sounds, etc.). Wikidata uses an elaborate and complex data model, which supports good-quality data but may also require knowledge about the data model for its effective use. If a high demand of human resources is required, it may represent an obstacle for cultural heritage data aggregators that are already operating with limited resources, and searching for more efficient ways to perform their data aggregation needs and remain sustainable.

In our study, we tested the hypotheses that Wikidata can be used efficiently, and with leading to quality-data, by machine-based methods operating with the data technologies and practices of the Semantic Web and linked data.

We follow, in Section 2, by describing related work on data aggregation based on linked data in cultural heritage. Section 3 presents the linked data crawling use cases that guide us in the design of the study of Wikidata, in light of our hypothesis. The setup and workflow of the study are presented in Section 4. Section 5 describes our linked data crawling software's architecture, and its relevant functionality for this particular study. Section 6 presents the results and our analysis. Section 7 concludes by summarizing the conclusions of the study and presenting future work.

2 Related Work

Linked data has a large diversity of research topics related to our work. Scalability is one of the most addressed topics, with many facets such as indexing, federated querying, and aggregation. The reuse of published linked data by third parties has revealed data quality to be a challenge as well, at the level of semantics and at the level of syntax [2, 3, 4]. Significant work has been done to facilitate the reuse of linked data by aggregation and data cleaning [5, 6]. Reasoning on linked data is also an active research topic, and a comprehensive analysis and description of techniques has been published [7].

Regarding cultural heritage, although the use of linked data has been the focus of much research, most of published literature addresses mainly the aspect of the publication of linked data [8, 9, 10] and do not fully address how the aggregation approach of cultural heritage can be based on the existing published cultural heritage linked data.

[1] https://europeana.eu

The most similar work to ours is that of the Dutch Digital Heritage Network[2] (NDE) and the Research and Education Space project[3] (RES). NDE is a Dutch national level program aiming to increase the social value of the collections maintained by the libraries, archives and museums in the Netherlands. NDE is still an ongoing project, and its initial proposals are based on specific APIs to enable data providers to centrally register the linked data URIs of their resources [11]. The current proposal of NDE, by being based in its own defined API, does not yet provide a solution purely based on linked data.

The Research and Education Space project finished in 2017, and has successfully aggregated a considerable number of linked data resources from cultural heritage sources. The resulting aggregated dataset can be accessed online[4], but an evaluation of its aggregation procedures and results was not published.

Generic technical solutions have been proposed by others for enabling aggregation of linked data (for example [12]). However, a standards-based approach has not yet been put into practice within cultural heritage.

The work presented in this paper is done in the context of the research activities, being carried out within the Europeana Network[5], for improving the network's efficiency and sustainability. Linked data has been identified in our past work as one of the technical solutions with application potential [13]. The work described in this paper is the continuation of a series of experiments addressing several Internet technologies for this purpose [14], particularly on the evaluation of Schema.org for cultural heritage [15] and linked data case studies [16].

3 Use cases

Our study aims to evaluate the usability of Wikidata as a data source for robots operating on the web of data, according to specifications and practices of linked data and the semantic web. In particular, we evaluate the possibilities for machine ontology-based reasoning on the RDF output of Wikidata, for RDF-crawling processes by robots and for using its data in other general data applications.

Our study addresses two use cases, which are illustrated in Fig. 1. The use cases represent different applications that a data consumer may use for processing data at different levels of semantic detail.

- Semantic Web general application – this type of application processes data solely based on the bottom and middle layers of the Semantic Web Stack [17]. These applications use general technologies, especially "meta-languages" such as RDF, RDF-Schema, OWL and linked data for crawling linked data. These applications

[2] https://github.com/netwerk-digitaal-erfgoed/
[3] https://bbcarchdev.github.io/res/
[4] https://bbcarchdev.github.io/res/collections
[5] The Europeana Network is a community of 1,700 experts with the shared mission to expand and improve access to Europe's digital cultural heritage, in the organization they work for and/or by contributing to shape Europeana's services.

require the use of properties that enable the crawling of linked data and the ontologies in use.
- Cross-domain semantic application – in addition to the functionality of Semantic Web general applications, this type of application processes data with additional requirements on the semantics of the data. They rely on general purpose, widely shared ontologies like the Schema.org vocabulary and SKOS (Simple Knowledge Organization System). These applications provide functionality that can be applied across different domains. They must either obtain the data that already use the ontologies they can consume (like Schema.org), or to find ontology alignments between these ontologies and the classes and properties used in the data. They thus require the ontologies used in the data source to be crawlable and to contain ontology alignments expressed in SKOS or OWL.

As mentioned in the introduction, these use cases are motivated by a third one, domain-specific semantic application, with data aggregation for Europeana being our core focus. This type of applications, while processing data like cross-domain applications, make further use of the semantics of the data. They indeed provide functionality for more specific purposes of a domain that requires a more detailed interpretation of the semantics of the data. They require the ability to convert the data to their domain-specific data model. Europeana uses the Europeana Data Model (EDM), which is its technological solution for data exchange with data providers. EDM is defined collaboratively withal the sectors represented in Europeana. Typically, they rely on ontology alignments defined by domain experts, whose work is enabled, or even only possible (regarding the amount of efforts these actors can afford), if widely used ontologies are used in the data.

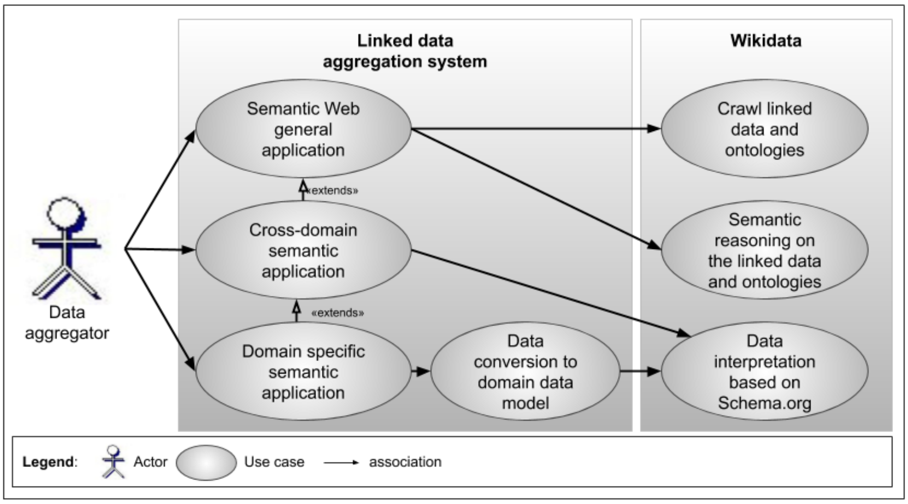

Fig. 1. Use cases of linked data consumption addressed in this study.

In this article, we present the results of our evaluation of the first two use cases. This work is the first milestone of our longer-term research interest to evaluate the use case of domain-specific applications. Note, however, that we include descriptions of the tasks involved in the third case at several points of this article, since it provides the complete context of our study.

4 Experimental setup

To assess the usability of Wikidata as a data source for our use cases, we have developed and applied software that includes components for linked data crawling, reasoning on RDF data and ontologies, data processing and data analysis. This software represents an extra iteration over previous work done for earlier experiments with linked data crawling [16]. As data sources for our study, we used the Wikidata Ontology through its linked data publication. As we aim at eventually studying the usability of Wikidata for the Europeana case, we focus on Wikidata resources that correspond to the Europeana dataset of cultural heritage resources.
The general overview of the study is as follows:

- Cultural heritage objects that are described in both Europeana and Wikidata were identified. The corresponding subsets were harvested for supporting our study.
- The Wikidata sample was converted to Schema.org. The conversion was based on the ontology alignments to Schema.org that exist in the RDF of Wikidata properties and classes. Wikidata classes were also converted, when present as 'object' in RDF triples. The RDF data regarding Wikidata's classes contains alignments to equivalent Schema.org classes. When an alignment with Schema.org was not found, an equivalence was searched in a more generic property/class by crawling up the hierarchy of Wikidata's classes and properties.

In future work, the data from Wikidata, after conversion to Schema.org, will be converted to our EDM data model, by applying mappings that we have defined in previous work, which studied the usage of Schema.org for describing cultural heritage resources [15].

A diagram of the setup for our experiments with Wikidata is represented in Fig. 2. All steps necessary for the complete study are shown, but at this stage, only part of setup was executed. The diagram illustrates the data sources, APIs, software components, samples, dataflows, and manual tasks. Our workflow is the following:

- Cultural heritage objects that are described in Wikidata were identified by querying the SPARQL API of Wikidata, and checking for Wikidata entities containing the property *Europeana ID*[6].
- The sample from Wikidata was collected using our software for linked data crawling, based on the set of Wikidata URIs identified in the previous step. The result was a dataset of 11.798 Wikidata entities about cultural objects. We actually identi-

[6] https://www.wikidata.org/wiki/Property:P727

fied 77.103 Wikidata entities containing the *Europeana ID*, however, during the course of our study, we identified that the values were invalid or obsolete in 65.305 of the cases, and we removed those entities from the subset of Wikidata we used.

- The values of the *Europeana ID* property were converted to their corresponding URIs at Europeana's linked data, by a simple process of prefixing the property value (a uniform URI structure is used by Europeana).
- The corresponding subset of Europeana was collected, also by our linked data crawler, based on the set of Europeana URIs. In future work, the Europeana subset will used for performing an evaluation of the data quality obtained from Wikidata.
- The Wikidata sample was further processed by conversion to Schema.org using our RDF converter software. At this stage of the workflow, we made our evaluation of the use cases for Semantic Web and cross-domain applications, and prepared reports about the RDF data, and about difficulties found during the crawling of Wikidata and the automatic interpretation of its RDF data.

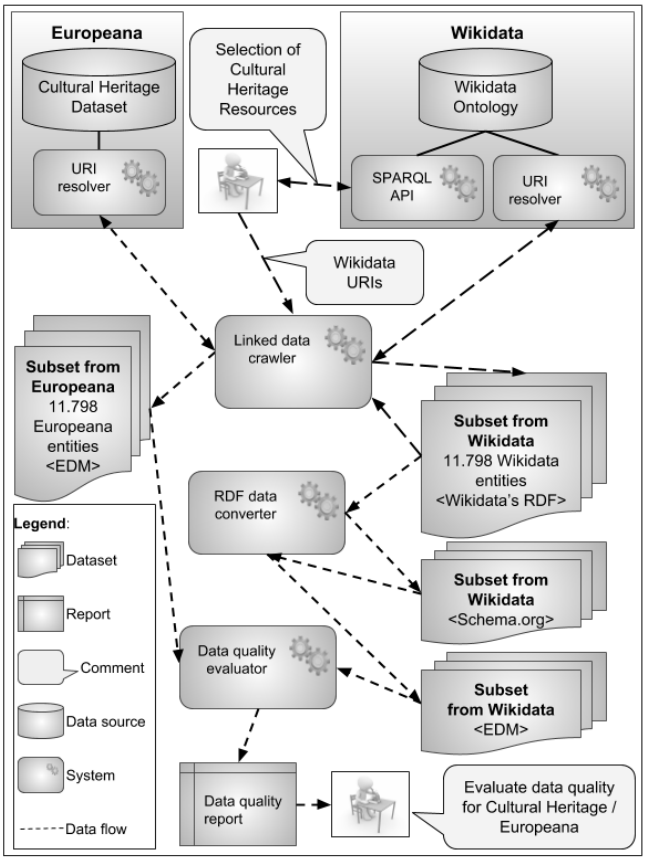

Fig. 2. The experimental setup

The remaining tasks of the workflow will be part of our future work. The Schema.org sample will then be converted to EDM, using the same software for conversion of RDF data. At this point, we will have two comparable subsets from both sources: Wikidata's subset represented in EDM RDF; and Europeana's subset, directly obtained in EDM RDF. To measure the level of quality of the data, we will apply another software component, which implements a metric for measuring the completeness of the data in EDM. The software generates a report that will support our final evaluation of the cultural heritage use case.

5 Architecture of the system for linked data aggregation

The architecture of the system supporting the execution of our experiment is illustrated in Fig. 3. The whole system is composed of 5 subsystems, supported by data repositories and resources that support its functions. The subsystems are the following:

- Workflow engine – This system allows a human operator to coordinate and monitor the execution of the workflow for the experiment.
- Linked data crawler – This system uses an implementation of the HTTP protocol, and implements the specifications and best practices of linked data, such as content negotiation and robots.txt files. It implements the interpretation of definitions of the members of linked data datasets according to the guidelines of Europeana[7], thus supports the interpretation of relevant properties for crawling linked data. The Europeana guidelines allow the use of properties from VoID, DCAT and Schema.org, for this purpose.
- Linked data interpreter – This system is capable of reasoning on properties for the purposes of interpretation of the semantics of data models' structural properties and URIs. It is based on properties from RDF, RDFS and OWL, to be able to interpret the concepts of *equality*, *types*, *class hierarchy and inheritance*, and *property hierarchy and inheritance*. It uses the definition of the data to populate a repository of statements (i.e. a triple store) to align data models. This repository may also be populated with statements that an operator of the system adds manually.
- RDF data converter – This system transforms an RDF graph into another RDF graph by executing data conversion operations. This data converter is specialized in RDF data and supports the underlying concepts of RDF data modeling. It converts RDF graphs based on specifications of detailed mappings between models. These specifications are represented in a machine-actionable form that allow the data converter to create detailed reports about the conversion of a dataset. These reports support data modelers in the definition of the mappings, and in the identification of possible improvements of data quality (e.g. missing data, invalid data types, missing mappings, etc) [16].

[7] https://github.com/nfreire/Open-Data-Acquisition-Framework/blob/master/opaf-documentation/SpecifyingLodDatasetForEuropeana.md

- Data quality evaluator –This system is a result of our research on mathematical metrics for evaluation of data quality in cultural heritage [18]. This system provides functionality for an aggregator to evaluate the potential of a dataset for its own purposes. It is currently being implemented and will be applied in this experiment. The system functions based on plugins that implement particular metrics or validations, and generates informative reports for the aggregator. In this study with Wikidata and Europeana, we will implement plugins for validation of EDM, and metrics of data completeness, whose definition has been in discussion within the Europeana Data Quality Committee[8].

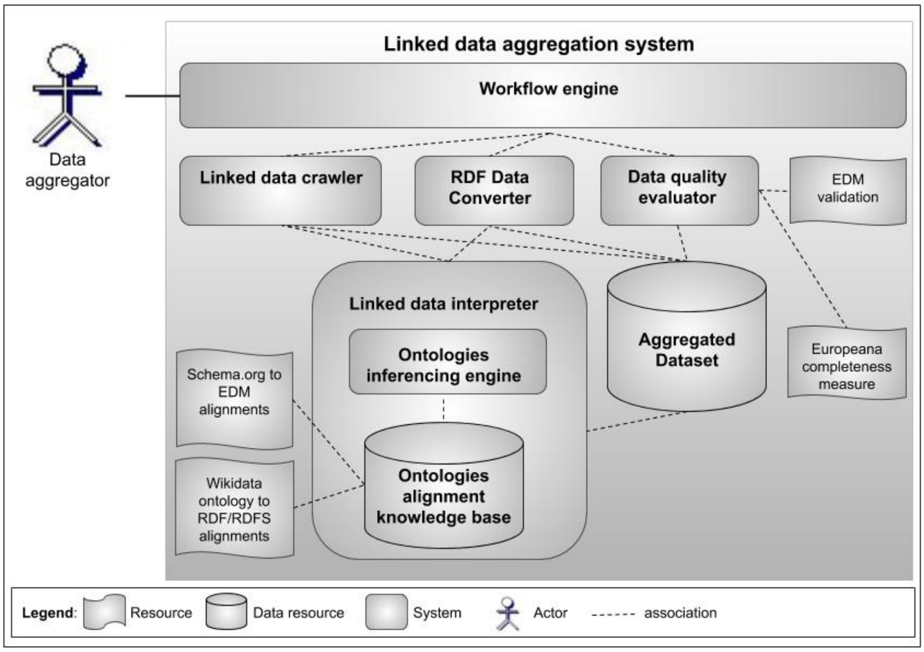

Fig. 3. High level system architecture of the linked data aggregation system

6 Results

In this section, we present the results of the study. It is important to note that the results of our study were gathered and checked for the last time on 12 of April 2019.

We observed that Wikidata's RDF presents some difficulties for cross-domain applications to aggregate and use the data. The difficulties are due to Wikidata's RDF using a very limited number of general data processing properties. Most of the properties in use, are labels, that are mostly useful for human users.

Wikidata has chosen to use properties from its own ontology instead of equivalent RDF, RDF-Schema, OWL or SKOS properties. Without some human intervention to

[8] https://pro.europeana.eu/project/data-quality-committee

support the application in interpreting Wikidata's properties, it would be impossible to use the data for any of our use cases. The count of occurrences of properties, in the collected dataset, of other namespaces than Wikidata's, are listed in Table 1. The main limiting aspect that blocks an advanced process in the interpretation of the data is the use of *rdf:type*[9]. It is used just to state that the RDF resource is an Item from the Wikibase ontology (*http://wikiba.se/ontology#Item)*, and for further types, the property *wdt:P31* is used. An advanced application that is able to search in all used properties' RDF resources, could bypass this limitation and interpret the data, but not in the case of Wikidata. This impossibility comes from the fact that Wikidata's RDF resources' URIs are not resolvable in all cases. In the case of property wdt:*P31*, it is stated in the data as http://www.wikidata.org/prop/direct/P31, which is not resolvable. The resolvable corresponding URI is http://www.wikidata.org/entity/P31.

Table 1. Properties, from non-Wikidata namespaces, in use in Wikidata's RDF output, of the 11.798 entities in our sample of cultural heritage records.

Property	Usage count
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	11.798
http://www.w3.org/2000/01/rdf-schema#label	29.571
http://www.w3.org/2004/02/skos/core#altLabel	2716
http://www.w3.org/2004/02/skos/core#prefLabel	29.571
http://schema.org/name	29.571
http://schema.org/description	64.563

We manually added the essential alignment statements in our knowledge base, so that we could proceed with the evaluation for the use cases that require the semantics of the data. Table 2 presents the alignment statements we added. In fact, most of the alignments that are required (6 out of 8) are already recorded in Wikidata, but they are expressed using predicates from the Wikidata namespace inhibiting the interpretation of their meaning by machines.

To evaluate the possibility of acquiring Schema.org semantics from Wikidata, we have used the equivalence relations that are stated in the RDF of the classes and properties' in Wikidata. When a resource does not state an equivalence with Schema.org, our linked data interpreter navigates the ontology's class/property hierarchy searching for an equivalence. If one is found the interpreter assumes it by inheritance.

To perform this evaluation for cross-domain applications, we faced the same difficulty with Wikidata's RDF – the URI's of properties are not resolvable. It was only

[9] For readability purposes, in this text we abbreviate namespaces as follows: *rdf* for *http://www.w3.org/1999/02/22-rdf-syntax-ns#*; *rdfs* for *http://www.w3.org/2000/01/rdf-schema#; owl* for *http://www.w3.org/2002/07/owl#*; *skos* for *http://www.w3.org/2004/02/skos/core#*; *schema* for *http://schema.org/*; *wikibase* for *http://wikiba.se/ontology#*; *wd* for *http://www.wikidata.org/entity/; wdt* for *http://www.wikidata.org/prop/direct/* .

possible for us to continue the study by manually adapting the code of the linked data interpreter to convert the URI's to the ones that Wikidata is able to resolve.

Table 2. Ontology alignments that were necessary for interpretation of semantics in Wikidata.

Aligned property	Wikidata property	Alignment existing in Wikidata (as wdt:P1628 or wdt:P1709)
rdf:type	wdt:P31(instance of)	*rdf:type*
rdf:Property	wdt:Q18616576 (instance of)	-
rdfs:Class	wdt:Q32753077 (instance of)	-
owl:sameAs	wdt:P2888 (exact match)	*skos:exactMatch*
owl:equivalentClass	wdt:P1709 (equivalent class)	*owl:equivalentClass*
owl:equivalentProperty	wdt:P1628 (equivalent property)	*owl:equivalentProperty*
rdfs:subClassOf	wdt:P279 (subclass of)	*rdfs:subClassOf*
rdfs:subPropertyOf	wdt:P1647 (subproperty of)	*rdfs:subPropertyOf*

An additional difficulty we have found, regards the interpretation of Wikidata's class and property hierarchy. It results from Wikidata using its own properties to state the class and property hierarchy. As in the first use case, we had to manually inject the essential alignment statements, which enable the interpretation of Wikidata's class and property structures (the alignments are listed in Table 2). Obtaining the equivalences between the classes and properties from Wikidata with Schema.org, faced the same difficulty – a Wikidata property is used to state the equivalences.

Table 3 presents statistics about the ontology alignments to Schema.org we found in Wikidata, considering only the classes and properties present in our subset. The listing of the individual alignments found may be consulted online[10]. In general, we found alignments for around 50% of the data elements in the sample. which, in our opinion, is a good indicative that many applications would be able to make use of the structured data. Particularly regarding classes, we found 102 distinct ones in use in the sample, 57% of which had alignments to Schema.org – 49% are direct alignments and 7,9% are alignments inherited from super classes. Regarding properties, we found 266 distinct ones in use in the sample, 44% of which had alignments to Schema.org – only direct alignments were found for properties.

Table 3. Statistics of the ontology alignments to Schema.org found in Wikidata, for the classes and properties in use in Wikidata's cultural heritage subset used in this study[9].

Type of data elements	Number of distinct data elements in Wikidata sample	Existing alignments	Existing alignments by inheritance	Total alignments
Classes	102	50 (49%)	7 (7,9%)	57 (55,9%)
Properties	266	44 (16,5%)	0 (0%)	44 (16,5%)

[10] https://github.com/nfreire/data-aggregation-lab/blob/master/data-aggregation-casestudies/documentation/wikidata/SchemaOrg-ontology-alignments-listing.md

7 Conclusion and future work

Currently, a human operator must assist linked data applications to interpret Wikidata's RDF, thus it requires training on data model behind Wikidata and its expression in RDF.

Our assessment is that Wikidata is only a few steps away from high quality machine interpretation, since there exists, in Wikidata, enough alignment data to RDF, RDFS, OWL, SKOS and Schema.org. Unfortunately, the semantics of the RDF output of Wikidata is *locked-in* the usage of predicates from Wikidata's own ontology, making them uninterpretable for data crawlers based on of properties for general data processing that the Semantic Web relies on.

The second difficulty is the use of namespaces that are not resolvable for Wikidata's properties. The reasons for this use of namespaces are documented [19], and justified as a way to represent characteristics of the predicates or objects of the triples. This practice is not standard, however, and other standard options are available that could be applied to address the reasons behind it.

We will continue our evaluation of Wikidata in our future work, by evaluating the use case where detailed semantics is required, using the specific domain of the cultural heritage network of Europeana, as we briefly described throughout this article. In later work, we expect to perform similar studies and this one, but on linked data published by data providers from the Europeana network.

References

1. Niggermann, E., Cousins, J., Sanderhoff, M.: Europeana Business Plan 2018 'Democratizing culture'. Europeana Foundation. (2018). https://pro.europeana.eu/files/Europeana_Professional/Publications/Europeana_Business_Plan_2018.pdf
2. Rietveld, L.: Publishing and Consuming Linked Data: Optimizing for the Unknown. In: Studies on the Semantic Web, vol. 21. IOS Press. (2016).
3. Radulovic, F., Mihindukulasooriya, N., García-Castro, R., Gomez-Pérez, A.: A comprehensive quality model for Linked Data. In: Semantic Web, vol 9, n. 1/2018. IOS Press. (2018).
4. Beek W., Rietveld L., Ilievski F., Schlobach S.: LOD Lab: Scalable Linked Data Processing. In: Pan J. et al. (eds) Reasoning Web: Logical Foundation of Knowledge Graph Construction and Query Answering. Reasoning Web 2016. Lecture Notes in Computer Science, vol 9885. Springer, Cham. (2017). doi: 0.1007/978-3-319-49493-7_4
5. Beek, W., Rietveld, L., Schlobach, S., van Harmelen, F.: LOD Laundromat: Why the Semantic Web Needs Centralization (Even If We Don't Like It). In: IEEE Internet Computing, vol. 20, issue 2. IEEE. (2016).
6. Fernández J.D., Beek W., Martínez-Prieto M.A., Arias M.: LOD-a-lot. In: d'Amato C. et al. (eds) The Semantic Web – ISWC 2017. ISWC 2017. Lecture Notes in Computer Science, vol 10588. Springer, Cham. (2017).
7. Hogan, A.: Reasoning Techniques for the Web of Data. In: Studies on the Semantic Web, vol. 19. IOS Press. (2014).

8. Simou, N.,Chortaras, A., Stamou, G., Kollias, S:. Enriching and Publishing Cultural Heritage as Linked Open Data. In: Ioannides M., Magnenat-Thalmann N., Papagiannakis G. (eds) Mixed Reality and Gamification for Cultural Heritage pp 201-223. Springer, Cham. (2017). doi:10.1007/978-3-319-49607-8_7
9. Hyvönen, E. Publishing and Using Cultural Heritage Linked Data on the Semantic Web. In: Ding, Y., Groth, P. (eds), Synthesis Lectures on the Semantic Web: Theory and Technology. (2012). doi: 10.2200/S00452ED1V01Y201210WBE003
10. Jones, E., Seikel, M. (eds). Linked Data for Cultural Heritage. Facet Publishing. (2016).
11. Meijer, E., Valk, S.: A distributed network of heritage information. White paper. (2017). Available online: https://github.com/netwerk-digitaal-erfgoed/general-documentation/blob/master/Whitepaper%20A%20distributed%20network%20of%20heritage%20information.md
12. Rietveld L., Verborgh R., Beek W., Vander Sande M., Schlobach S.: Linked Data-as-a-Service: The Semantic Web Redeployed. In: Gandon F., Sabou M., Sack H., d'Amato C., Cudré-Mauroux P., Zimmermann A. (eds) The Semantic Web. Latest Advances and New Domains. ESWC 2015. Lecture Notes in Computer Science, vol 9088. Springer, Cham. (2015).
13. N. Freire, H. Manguinhas, A. Isaac, G. Robson, J. B. Howard, "Web technologies: a survey of their applicability to metadata aggregation in cultural heritage.", in: L. Chan, F. Loizides (eds.), Information Services & Use Journal, volume 37, issue 4, Expanding Perspectives on Open Science: Communities, Cultures and Diversity in Concepts and Practices. (2018).
14. Freire, N., Robson, G., Howard, J. B., Manguinhas, H., Isaac, A. Metadata Aggregation: Assessing the Application of IIIF and Sitemaps within Cultural Heritage. In: 21st International Conference on Theory and Practice in Digital Libraries. (2017).
15. Freire, N., Charles, V., Isaac, A.: Evaluation of Schema.org for Aggregation of Cultural Heritage Metadata. In: Gangemi A. et al. (eds) The Semantic Web. ESWC 2018. Lecture Notes in Computer Science, vol 10843. Springer, Cham. (2018). doi:10.1007/978-3-319-93417-4_15.
16. Freire, N., Meijers, E., Voorburg, R., Cornelissen, R., Isaac, A., de Valk, S.: Aggregation of Linked Data: A case study in the cultural heritage domain. 2018 IEEE International Conference on Big Data (Big Data), IEEE. (2018).
17. Curé, O., Blin, G. (eds): Chapter Three - RDF and the Semantic Web Stack. In RDF Database Systems, pp 41-80. Morgan Kaufmann. (2015). doi: 10.1016/B978-0-12-799957-9.00003-1.
18. P. Király, P., Stiller, J., Charles, V., Bailer, W, Freire, N.: Evaluating Data Quality in Europeana: Metrics for Multilinguality. In: Garoufallou E., Sartori F., Siatri R., Zervas M. (eds) Metadata and Semantic Research. MTSR 2018. Communications in Computer and Information Science, vol 846. Springer, Cham. (2018). doi: 10.1007/978-3-030-14401-2_19.
19. Erxleben, F., Günther, M., Krötzsch, M., Mendez, J., Vrandečić, D.: Introducing Wikidata to the Linked Data Web. In Proceedings of the 13th International Semantic Web Conference, pp. 50–65. Springer, Cham. (2014).