



# Cultural heritage on the Semantic Web: The Europeana Data Model

Ana Luísa Silva 

Centre for Interdisciplinary Studies, University of Coimbra, Portugal

Ana Lúcia Terra 

Centre for Interdisciplinary Studies, University of Coimbra, Portugal

International Federation of  
Library Associations and Institutions  
1–15

© The Author(s) 2023



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/03400352231202506

[journals.sagepub.com/home/iff](https://journals.sagepub.com/home/iff)



## Abstract

The Semantic Web allows data to be linked on the Web and structures information for use by humans and machines. Furthermore, it makes relationships between data explicit, enabling the creation of Linked Data. Based on a literature review, the principles and technologies underlying Linked Data are presented – namely, the Resource Description Framework and models developed for libraries, archives and museums. Europeana aggregates European institutions' digital cultural heritage, having developed a model that follows linked-data principles. For a deeper understanding of this model, the Europeana Data Model is presented with examples of two representation approaches and the advantages of metadata enrichment in information discovery. The experience of the University of Coimbra with Europeana is briefly explained. Finally, the authors discuss the challenges that cultural heritage institutions face in adopting these models and freeing their information from the silos it is in, taking advantage of the potential that Linked Data provides.

## Keywords

Europeana Data Model, libraries, archives, museums, convergence, Linked Data, Resource Description Framework, Semantic Web

## Introduction

The Semantic Web is an evolution of the World Wide Web (Web) that has as its main purpose the interconnection of data through a set of tools and techniques to structure and relate information on the Web so that it can be shared, discovered, integrated and reused efficiently by both humans and machines. To achieve this goal, information on the Web must be available in a standardized, searchable format that can be managed by Semantic Web tools. In addition to representing data in a structured way, it is essential to make the relationships between data explicit, which allows for the creation of a vast collection of interrelated data sets, called Linked Data. If the information is freely accessible, the term Linked Open Data is used.<sup>1</sup> Linked Data is based on the Resource Description Framework (RDF) semantic metadata model.

The need to evolve the traditional metadata formats used in libraries, museums and archives into Linked Data models has triggered initiatives in several

institutions, leading to the emergence of different conceptual models for each of these areas. One of these models – the RDF-based Europeana Data Model – was developed to support Europeana, a digital platform that aggregates all kinds of digital content from cultural heritage institutions across Europe.<sup>2</sup>

This article presents the technologies associated with Linked Data and describes the models developed for libraries, museums and archives, as well as the Europeana Data Model developed for Europeana. It aims to prove the applicability of Linked Data technologies in structuring and enriching the metadata of digital objects in libraries, archives and museums, and to highlight the advantages of implementations based

---

### Corresponding author:

Ana Luísa Silva, CEIS20 – Centre for Interdisciplinary Studies, University of Coimbra, Rua Augusto Filipe Simões 33, 3000-457 Coimbra, Portugal.  
Email: [als@uc.pt](mailto:als@uc.pt)

on Linked Data models in the interoperability, discovery and visualization of the information of digital resources in cultural heritage institutions. Some possibilities for future work are also presented.

This work has been developed from a literature review, considering the reference documentation of the World Wide Web Consortium (W3C),<sup>3</sup> the documentation on the Europeana Data Model and Europeana's technical recommendations, and scientific articles mainly from the last five years.

## Linked Data

Much of the information we get on the Web is provided through HTML (HyperText Markup Language) pages, and by using hyperlinks we can scroll through linked pages.<sup>4</sup> However, if we perform Web searches on a particular subject and we get several links in the results page, we need to open each one to understand its relevance to the search we are doing. The Semantic Web allows for the evolution of information discovery mechanisms to a semantic level through the structured representation of data and its connections, so that data can be processed by machines. Linked Data is the core of the Semantic Web and follows the principles defined by Tim Berners-Lee (2006): (1) use Uniform Resource Identifiers (URIs) as persistent data identifiers; (2) use HyperText Transfer Protocol (HTTP) URIs so that they can be accessed; (3) use standardized vocabularies, which allow machines to interpret data in the same way and ensure interoperability and the ability to interrogate multiple repositories; and (4) include links to other URIs so that data can be connected.<sup>5</sup>

Linked Data Web technologies are part of a set of W3C recommendations, enabling the creation of Web data repositories, building vocabularies and writing rules to manipulate data. The underlying technologies for Linked Data are the RDF; various serialized representations of the RDF (RDF Extensible Markup Language (RDF/XML), Notation3 for RDF (N3), Turtle, N-Triples, RDF/JSON);<sup>6</sup> data query language (SPARQL);<sup>7</sup> and the definition of vocabularies and ontologies (the RDF Schema, the Simple Knowledge Organization System (SKOS), Web Ontology Language (OWL)) (World Wide Web Consortium, 2015).<sup>8</sup>

### Resource Description Framework

The RDF is a model for describing resources and the relationships between them in a flexible and domain-independent way. Its basic structure is a three-element tuple (triple), subject  $\rightarrow$  predicate  $\rightarrow$  object, which allows sentences about resources to be made. An example is shown in Figure 1.

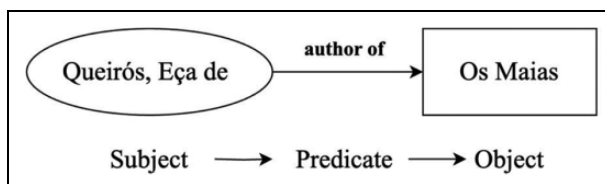


Figure 1. Example of an RDF triple.

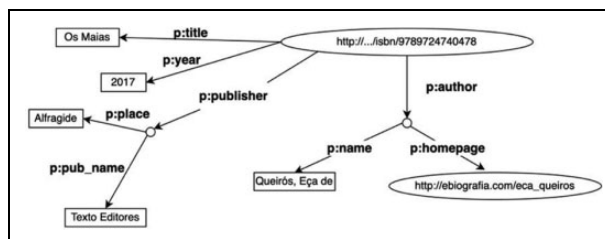


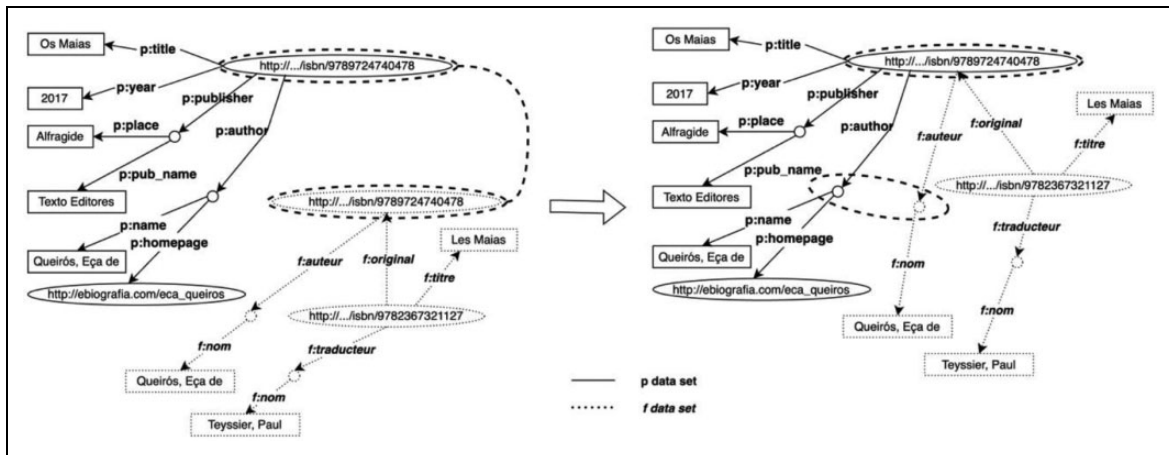
Figure 2. Data representation (data set p).

For the example in Figure 1, we can write the triple as follows: ( $\langle$ http://.../queiros $\rangle$ ,  $\langle$ http://.../author $\rangle$ ,  $\langle$ Os Maias $\rangle$ ). The subject and predicate elements are Web resources, with a corresponding URI accessible via the HTTP protocol. The subject and object elements can have a literal value or a URI. The predicate element establishes a relationship or property between the subject and the object.

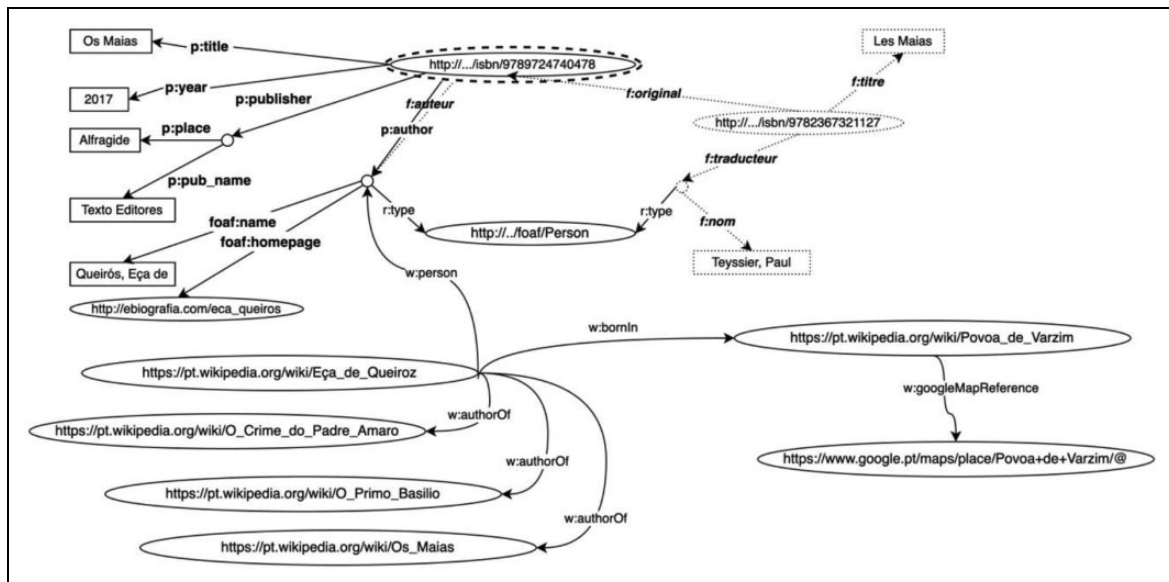
Multiple triples can be combined, building a network of unlimited information. For the triple shown in Figure 1, if we replace the literal value of the object with a URI, we get a labelled graph and, with multiple triples, we can represent more information (Figure 2) from a database, which is labelled 'p'.

A similar data set corresponding to a translation of the same work, database f, is shown in Figure 3. The existence of a common URI between the two structures makes it possible to link the two graphs and automatically obtain information about the original title of the work from database f, which is not in that database. One of the great advantages of Linked Data is that these databases can exist in different places on the Web.

We can improve the merged result by adding some information considering the equivalence of the predicate naming: 'p: author' is the same as 'f: auteur'. Moreover, both correspond to a resource category already defined in the community, such as 'Person' ('foaf: Person')<sup>9</sup>, with information about the name, date of birth and so on. The new representation of information, resulting from the addition of new phrases and existing terminology, allows the f-database user to ask questions such as 'What is the web page of the author of the book?', even if this information is not in their database. By connecting to an already defined data set ('Person'), it is possible to combine this data set with other data sources, such as Wikipedia.<sup>10</sup> The resulting graph is shown in Figure 4.



**Figure 3.** Evolution of RDF representation with additional information.  
Source: Adapted from Herman (2010).



**Figure 4.** RDF representation with additional information and links to common vocabularies (Person).  
Source: Adapted from Herman (2010).

```

1      <rdf:RDF
2          xmlns:f="http://www.editeur.fr"
3          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4          <rdf:Description rdf:about="http://.../isbn/9782367321127">
5              <f:original rdf:resource="http://.../isbn/9781501146213"/>
6              <f:titre xml:lang="fr">Les Maias</f:titre>
7          </rdf:Description>
8      </rdf:RDF>

```

**Figure 5.** Example of RDF/XML serialization.  
Source: Adapted from Herman (2010).

To present RDF data, several serialization formats can be used, including RDF/XML, Turtle, N-Triples, Trig, N-Quads, JSON-LD and RDFa. Figure 5 shows an example of an RDF and its serialization in RDF/XML.

### RDF vocabularies, ontologies and SPARQL

To allow interoperability, the RDF representation of data must be meaningful and contextualized. The RDF Schema is an extension of the RDF that allows RDF vocabularies to be defined – that is, it provides

```

<rdfs:Class rdf:about="&person;Person">
  <rdfs:label xml:lang="en">Person</rdfs:label>
  <rdfs:comment xml:lang="en">An individual person who may be dead or alive, ...</rdfs:comment>
  <rdfs:isDefinedBy rdf:resource="&person;"/>
  <rdfs:subClassOf rdf:resource="&foaf;Person" />
  <rdfs:subClassOf rdf:resource="&schema;Person" />
  <dcterms:identifier>person:Person</dcterms:identifier>
</rdfs:Class>

<rdf:Property rdf:about="&person;placeOfBirth">
  <rdfs:label xml:lang="en">place of birth</rdfs:label>
  <rdfs:comment xml:lang="en">The range of place of birth is dcterms:Location and schema:Place.</rdfs:comment>
  <rdfs:range rdf:resource="&schema;Place" />
  <rdfs:range rdf:resource="&dcterms;Location" />
  <rdfs:isDefinedBy rdf:resource="&person;"/>
  <dcterms:identifier>person:placeOfBirth</dcterms:identifier>
</rdf:Property>

```

**Figure 6.** Example of vocabulary (class and property).

Source: Archer and Perego (2012).

the mechanisms for describing groups of related resources and the terms used to relate those resources ('predicate' in the RDF triple). In a simplified description, it is equivalent to a database schema, but in this case the definition is distributed across the Web, can be defined by anyone and will become common usage of Linked Data if it is used by many people on the Web (Wood et al., 2014). RDF vocabularies are sets of terms to describe resources. Each term can be a class or a property. Classes bring together resources with the same properties. Properties can be relationships (between resources) or attributes (characteristics). An example is shown in Figure 6.

Vocabularies are defined in HTTP URIs and should follow the Linked Data principles, reuse existing vocabularies, and provide useful information on how the terms they define should be used. They usually have a prefix for the conceptual area they represent (namespace), which is commonly used to simplify RDF serializations. For example, the Dublin Core namespace is defined by the URI <http://purl.org/dc/elements/1.1/> and uses the prefix 'dc'.<sup>11</sup> The RDF itself is defined by the URI <http://www.w3.org/1999/02/22-rdf-syntax-ns#> and uses the prefix 'rdf:'. The lines in bold in Figure 5 show the URIs of the vocabularies used and their prefixes. In line 4 and line 6 of Figure 5, the formulations `rdf: about` and `f: titre` describe attributes of the resource defined by `rdf: Description`, whose identifier is given in `rdf: about="http://.../isbn/9782367321127"` (line 4). In line 5, the attribute `f: original` defines the relationship between the described resource and the one with the URI identified through the attribute `rdf: resource="http://.../isbn/9781501146213"`.

When different data sets use the same terms, vocabularies serve to resolve ambiguity or provide additional information that allows the discovery of new data relationships. Of great importance for the library

community is the SKOS, an RDF vocabulary that aims to transform traditional knowledge organization systems (thesauruses, classification systems, subject headings or taxonomies) into Linked Data structures, and is concerned with maintaining compatibility with ISO 25964-2 (International Organization for Standardization, 2023).

Sometimes, the RDF Schema is not enough to guarantee all possible requirements, and the support of ontologies on the Semantic Web is necessary. An ontology is a knowledge organization model that defines a set of concepts and the relationships between those concepts for a given domain. It supports automated reasoning and data inference using logical rules, and allows knowledge to be shared and reused by people and machines. There are two W3C recommendations regarding ontologies: OWL and the Rule Interchange Format (Raafat, 2017).

To obtain data from the Semantic Web and use it to build applications or integrate it into web portals or web pages, it is necessary to have a language for searching and retrieving information: this is SPARQL, short for "SPARQL Protocol and RDF Query Language". SPARQL can search local RDF (serialization) files or RDF databases anywhere on the Web, provided that these databases have a SPARQL endpoint service.<sup>12</sup> The advantage of SPARQL and Linked Data is the possibility of combining local and remote queries to obtain the desired data (Wang and Yang, 2018).

### Linked Data in libraries, archives and museums

Libraries, archives, museums and other cultural heritage organizations are constantly seeking to improve the visibility and scope of their collections to serve diverse audiences, from academia to the general public. To this end, data is shared on cultural heritage aggregator portals (national or international), such

as Europeana and the Digital Public Library of America, or presented on general platforms, such as DBpedia or Wikidata, following Semantic Web models.<sup>13</sup> This requires that the information is organized in well-defined structures, following standardized and semantically rigorous publication models (Zapounidou et al., 2017).

According to Joudrey and Taylor (2018: 129), ‘metadata is structured information that describes the important attributes of information resources for the purposes of identification, discovery, selection, use, access, and management’. They further state that, for metadata to fulfil these purposes, it needs to be interoperable, as well as flexible and expandable. Interoperability allows the interconnection of information between technologically distinct systems with minimal loss and must be present at the semantic, syntactic and structural levels. Semantic interoperability refers to how different metadata schemas assign meaning to their elements and the correspondence between those elements. Syntactic interoperability refers to data encoding formats (e.g. Encoded Archival Description, XML, machine-readable cataloguing (MARC)) and the ability to exchange and use data from other systems. Structural interoperability concerns the data model used (e.g. whether it follows the RDF structure) (Joudrey and Taylor, 2018).

Detailed models and schemas for describing information resources have long existed, but they are not compatible with Web standards, let alone the Semantic Web principles. Several libraries and standards organizations have developed conceptual models and metadata standards that are suitable for the Web and particularly for Linked Data environments.

### *Linked Data models in libraries*

For several years, initiatives have been developed to foster the use of Linked Data in the library field. The Library of Congress, IFLA and Online Computer Library Center (OCLC), among others, have been committed to these efforts.

The IFLA developed the IFLA Library Reference Model (IFLA LRM), a standard that resulted from the consolidation of previously developed models for bibliographic information (Functional Requirements for Bibliographic Records (FRBR), Functional Requirements for Authority Data, Functional Requirements for Subject Authority Data), which were designed to support and promote the use of bibliographic data in Linked Data environments in the library universe (Joudrey and Taylor, 2018). Addressing print materials in more detail, this standard does not apply to many existing museum resources or presents

difficulties for application in archival collections. This is one of the reasons that contributed to the development of specific conceptual models for these types of resources.

Based on the FRBR, Resource Description and Access was launched in 2011. It is structured for Linked Data applications and is a user-centred set of data elements, guidelines and instructions for creating metadata for library and cultural heritage resources. Resource Description and Access replaced the second edition of the Anglo-American Cataloging Rules (Park and Kipp, 2019).

Following this new model, the Library of Congress developed the Bibliographic Framework. This is an initiative to evolve bibliographic description standards towards a Linked Data model and to enable greater use of bibliographic information within and beyond the library community (Library of Congress, 2015). The Bibliographic Framework seeks to ensure the transition of all bibliographic heritage in MARC 21<sup>14</sup> to the Semantic Web through a Linked Data model based on the RDF. In its current version (2.0), it is based on Linked Data technologies (the RDF Schema, the SKOS and OWL) and includes two main components: the Bibliographic Framework data model, which describes the structure of bibliographic resources and their potential relationship with other structures, and the Metadata Authority Description Schema in RDF, for knowledge organization.<sup>15</sup> In addition to developing the Bibliographic Framework, the Library of Congress (n.d.) makes available its ontologies, controlled vocabularies and other bibliographic description lists for use in Linked Data.

OCLC is another organization that is involved in research and development projects in the field of Linked Data. Its WorldCat database aggregates millions of bibliographic records from libraries around the world. It adopted the Schema.org vocabulary, which includes entities, entity relationships and actions, and was developed by Google, Microsoft, Yahoo and Yandex.<sup>16</sup> This allows WorldCat to expose its data for easier use in search systems and other applications. OCLC led the work to reconcile the Schema.org vocabulary with the Bibliographic Framework, and participated in the development of a Schema.org extension for bibliographic information. Also developed and maintained by OCLC, the Virtual International Authority File (VIAF) is an international authority file that includes authority records from national libraries around the world and has been available as Linked Data since 2009 (Wang and Yang, 2018).<sup>17</sup>

### *Linked Data models in archives*

For archival information, the Expert Group on Archival Description (2021) of the International Council on Archives is developing a conceptual model for archival description, which merges four standards for archival documentation (the General International Standard for Archival Description; the International Standard Archival Authority Records—Corporate Bodies, Persons and Families; the International Standard Description of Functions; and the International Standard Description of Institutions with Archival Holdings) and is called the Records in Contexts—Conceptual Model. Like the IFLA LRM for bibliographic description, this model identifies entities, the properties of entities and the relationships between entities. Associated with the conceptual model, a specific implementation was developed using OWL: the Records in Context Ontology. This enables archival description using Linked Data techniques. Both formulations are part of the Records in Contexts standard, which also includes Records in Contexts—Application Guidelines and Records in Contexts—Introduction to Archival Description. The most recent version of the Records in Contexts—Conceptual Model (0.2) is from July 2021.

### *Linked Data models in museums*

The International Committee for Documentation (CIDOC) of the International Council of Museums has been working on a general data model – the CIDOC Conceptual Reference Model – to enable museum libraries, archives and other cultural heritage institutions to improve the sharing of information between each other (International Council of Museums, n.d.). It currently corresponds to ISO 21127:2014 – *Information and documentation – A reference ontology for the interchange of cultural heritage information* (International Organization for Standardization, 2020). According to Stein and Balandi (2019), there have been several attempts to provide a generally valid path for transferring data from Lightweight Information Describing Objects – CIDOC’s recommended XML schema for the delivery of museum metadata – to representations suitable for the Semantic Web.<sup>18</sup> The CIDOC Conceptual Reference Model is a complex system consisting of 94 entities (classes) and 168 relationships (properties) to represent and share cultural heritage metadata, and includes the representation of events associated with an object throughout its existence. Although its base structure is different from the models previously presented, its information can be mapped to other data models, such as the Dublin

Core, Encoded Archival Description and the FRBR (Joudrey and Taylor, 2018).

### *Other models*

In trying to align the CIDOC Conceptual Reference Model and IFLA FRBR and LRM, the new IFLA FRBRoo and LRMoo (object-oriented) models were developed (Bekiari et al., 2015). These include not only bibliographic entities, but also the activities related to their intellectual creation, editing and publication over time (Zapounidou et al., 2017). The Europeana Data Model, developed by Europeana, will be discussed in detail in the following section.

### **Case study: Europeana**

Europeana is a digital platform resulting from the European Union’s initiative to create a European virtual library to make Europe’s cultural heritage visible and digital repositories accessible. Created in 2008, managed by the Europeana Foundation, and integrating Europe’s major libraries, archives, museums, audiovisual archives and cultural institutions, it is a single access point to the digital cultural heritage of approximately 3700 European institutions in 25 languages.

As explained on its website: ‘The digital collections of Europeana provide multiple perspectives on historical, scientific and cultural developments across Europe and beyond’ (Europeana Foundation, 2008). To demonstrate the potential of this digital heritage, it offers mentoring programmes, online courses, pedagogical materials and teaching platforms. As further stated on its website:

Europeana empowers the cultural heritage sector in its digital transformation. We develop expertise, tools and policies to embrace digital change and encourage partnerships that foster innovation. We make it easier for people to use cultural heritage for education, research, creation and recreation. Our work contributes to an open, knowledgeable and creative society. (Europeana Foundation, 2008)

Europeana contains over 53 million digital objects in the public domain or with free access licenses in several languages, organized by various themes and geographical areas. It offers several services: research, thematic digital archives and collections, virtual exhibitions and a blog. To ensure the quality of its metadata, Europeana relies on a network of aggregation partners that collect data from the various institutions and verify and enrich it with additional information, such as georeferencing data or links to

other data on people, places or subjects (Petras et al., 2017). These aggregation partners can be national in scope, such as Portugal's Registo Nacional de Objetos Digitais (National Register of Digital Objects); in the cultural sector, such as Archives Portal Europe, which aggregates information on archival objects from European countries and national archives, or MUSEU-HUB, for museums; dedicated to certain types of material, such as Manuscriptorium or Musical Instruments Museums Online (MIMO); or thematic, such as the European Fashion Heritage Association or the Technische Informationsbibliothek portal for scientific videos.<sup>19</sup>

### *Europeana Data Model*

Europeana aggregates digital objects, including works of art, books, photographs, letters, maps, newspapers, music and videos about art, archaeology, fashion, science, sports and more. They come from many repositories and use a wide variety of descriptive metadata standards. To provide access to digital cultural heritage and make it user-friendly, it was necessary to find a way to integrate these standards and make the metadata of all the partner institutions interoperable. For this purpose, a set of metadata that could encompass all this content with minimal loss of information had to be defined. To this end, the Europeana Semantic Elements Model was developed; this data model was used when the prototype of Europeana was launched in November 2008 (Ciocoiu, 2018).

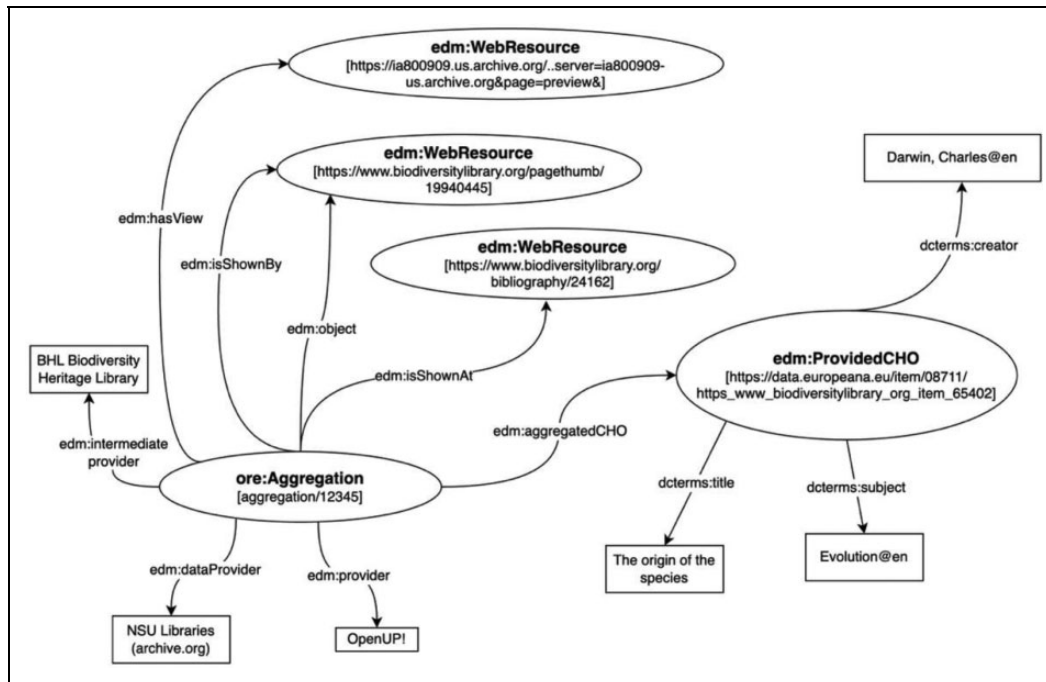
The Europeana Semantic Elements Model is a linear structure that uses the Dublin Core Metadata Element Set with some Europeana-specific extensions.<sup>20</sup> However, the weaknesses of the Europeana Semantic Elements Model were quickly identified – namely, the lack of distinction between the physical object in the institution's collection and its digital representation; poor expandability; little flexibility to deal with the specificities of some objects; and the difficulty of creating semantic links between objects and resources. It became essential for it to evolve into a more flexible data model – the Europeana Data Model (EDM) – which is capable of overcoming the Europeana Semantic Elements Model's problems and based on a Linked Open Data model. The Europeana Data Model is RDF-based and therefore an open and comprehensive model that allows links to other digital resources and data enrichment. Its development began in 2009 and is the result of the close collaborative work of Europeana experts in different fields with its internal team. The goal was to consider as many requirements as possible in the domains being worked on in an attempt to define a flexible model to

represent data from the different sectors of Europeana. The data model was consolidated with some partners, such as the Deutsche Digitale Bibliothek and the Digital Public Library of America, which adopted the Europeana Data Model as the basis for their data models (Petras et al., 2017).<sup>21</sup> The more recent infrastructure ROSSIO,<sup>22</sup> 'a free and open-access platform for aggregating, organizing, and connecting the digital resources in the Social Sciences, Arts and Humanities provided by Portuguese higher education and cultural institutions' (Silva et al., 2022: 1), has adopted a Europeana Data Model profile that is capable of representing digital objects from different domains defined by Portuguese academic and cultural heritage institutions.

The Europeana Data Model solves the problems of the Europeana Semantic Elements Model in various ways: it distinguishes between the objects provided and their digital representations; it distinguishes object records from metadata records; it allows the representation of complex objects and for the detailing and expanding of the metadata format;<sup>23</sup> it ensures compatibility with different levels of abstraction in the description; and it allows the use of contextual resources to incorporate concepts from controlled vocabularies.

**Data representation.** The Europeana Data Model defines its own vocabulary, containing specific elements (classes and properties), and reuses elements from other vocabularies such as the Dublin Core, Open Archives Initiative Object Reuse and Exchange, SKOS and Creative Commons. It also defines a set of descriptive and contextual properties that characterize the different aspects of a resource and relate it to other entities in its context. It can incorporate content with different description models (MARC 21, Encoded Archival Description, Lightweight Information Describing Objects) and with various levels of granularity. It also allows the same (physical) object to have several descriptions from different providers and to represent information added by Europeana.

The Europeana Data Model includes 11 specific classes and 6 classes belonging to other namespaces. For each class, it defines a set of properties, some of which are mandatory. In a simplified approach, a digital object provided to Europeana can be represented by three main classes and four contextual classes. Figure 7 is a graphical representation of a digital object from Europeana. The object is available at: [http://www.europeana.eu/pt/item/08711/https://www.biodiversitylibrary.org/item\\_65402](http://www.europeana.eu/pt/item/08711/https://www.biodiversitylibrary.org/item_65402); the original document is available at: <https://www.biodiversitylibrary.org/bibliography/24162>.



**Figure 7.** Digital object schema in the Europeana Data Model.  
Source: Adapted from Isaac and Clayphan (2013).

The class ‘Provided Cultural Heritage Objects’ (‘edm:ProvidedCHO’) refers to the original object (which can be a physical object (book, painting, photograph, map, drawing) or a digital original). All of the metadata describing the object must be part of this class. Each original object can have one or more accessible digital representations associated with it, some of which may be used as ‘thumbnails/previews’ and belong to the ‘Web Resources’ class (‘edm:WebResource’).

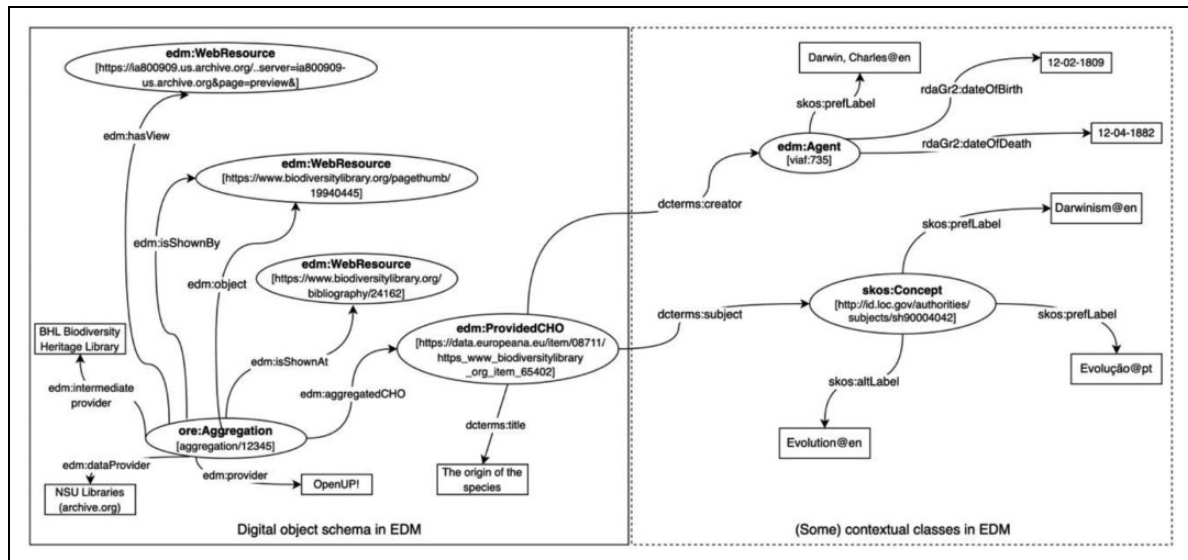
The class ‘ore: Aggregation’ links an original object to its digital representations described by a provider and included in Europeana, allowing the whole structure to be considered as a logical unit. Using the properties defined in the Europeana Data Model, each instance of the ‘ore: Aggregation’ class is related to a resource corresponding to the original object, ‘edm:ProvidedCHO’, through the ‘edm:aggregatedCHO’ property, and to one or more ‘edm:WebResource’ resources that are its digital representation(s). The instances of the ‘edm:WebResource’ class may have the following links (among others): ‘edm:isShownAt’, which refers to the full display of the object on the original website; ‘edm:isShownBy’, which refers to the main representation of the object that may be used to generate the thumbnail by the Europeana system; and ‘edm:object’, which also refers to the main representation of the object and normally has the same URL as ‘edm:isShownBy’. For subsequent representations, ‘edm:hasView’ should be used.

Also associated with the class ‘ore: Aggregation’ are three types of properties to identify the content providers: ‘edm:dataProvider’, which is mandatory, identifies the institution that is responsible for the original object and the information about that object; ‘edm:intermediateProvider’ is the intermediate organization that selects, collects or curates the data, and makes the information available in an aggregator; and ‘edm:provider’ is the provider from which Europeana obtains the information to load into its system.

Descriptive metadata can be attached to the original object, usually using the Dublin Core schema. Direct links are created between the described object and its characteristics, which can be expressed through simple text strings such as ‘dcterms: title’, ‘dcterms: subject’ or ‘dcterms: creator’ (see Figure 7). These links can also be made to other entities in other data sets. All of the entities have a unique identifier (URI) so that they can be referenced internally (Europeana Foundation, 2014).

**Context classes.** Some of the values of the descriptive metadata may be related to other entities, such as author, with more details associated with them. If this metadata is populated with URIs – that is, in accordance with the Linked Data principles – the system (Europeana) will create classes based on these URIs. To enable semantic enrichment, the Europeana Data Model provides context classes: ‘edm:Agent’ to represent people or organizations; ‘edm:Event’ for





**Figure 8.** Digital object schema in the Europeana Data Model with context classes for author and subject.

events; ‘edm:Place’ for geographical or location entities; ‘edm:TimeSpan’ for time periods or dates; and ‘skos: Concept’ for all the entities that comprise knowledge organization systems (thesauruses, classification systems, authority files, etc.). Figure 8 shows a new version representing the same model using context classes.

If we consider the information regarding the author of the book, in Figure 7, additional details about the author are not provided. This description can be improved by creating an explicit link (URI) between the book described and a resource that delivers such information – the Virtual International Authority File record on Darwin (<https://viaf.org/viaf/735/>) – by introducing the class ‘edm:Agent’ in the link ‘dcterms: creator’ (see Figure 8). This semantic enrichment of data brings huge benefits for the information search and discovery process. The contextual information added with the class ‘skos: Concept’ in the link ‘dcterms: subject’, like the value ‘Evolution’, leverages links to other entities and allows the use of all the information available in Wikipedia about each subject, including its translation into several languages.

**Descriptive metadata: object-centric and event-centric approaches.** The Europeana Data Model allows two approaches to descriptive metadata. In the object-centric approach, there is a direct link between the object and its characteristics, using literal values or links to other resources. This approach is the one used in Figures 7 and 8. The event-centric approach emphasizes the description of the different events in which the object was involved, providing the creation of a more complete network of entities that reflect the

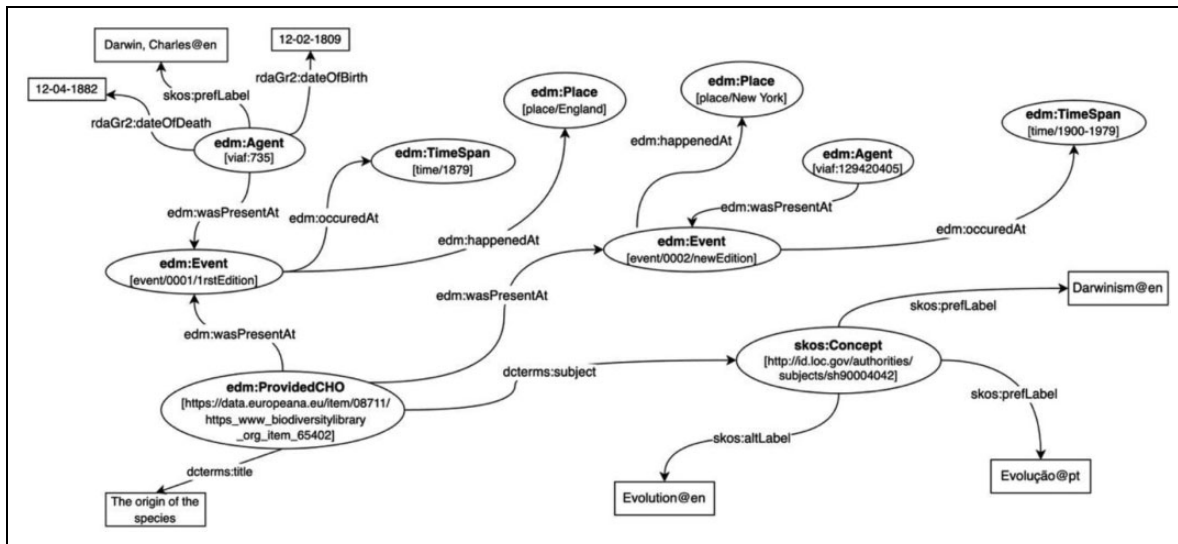
history of the object. The CIDOC Conceptual Reference Model, mentioned above, underlies this approach, which is adopted by some of the institutions participating in Europeana, particularly in the areas of museums and archaeology.

The ‘edm:Event’ class serves this purpose with three properties that relate to the resource being described: ‘edm:wasPresentAt’ – the event in which it was involved; ‘edm:happenedAt’ – the place where it happened; and ‘edm:occurredAt’ – the time when it occurred. To maintain consistency, Figure 9 shows a representation of the same digital object in an event-centric approach. Two events have been added – the publication of the first edition of the book (1879) and the new edition present in Europeana (1900-1979).

This approach can serve as a basis for digital storytelling.

**Hierarchical objects and other relationships.** Europeana has always considered the treatment of hierarchical objects an important topic, since various types of hierarchies can be found in the cultural heritage domain. These types of hierarchies include vertical relationships (whole–part) as well as horizontal relationships (sibling). The main types of relationships that appear in Europeana are hierarchical descriptions of objects from libraries (volumes, chapters, issues and articles), hierarchical objects in audiovisual archives, hierarchies for archaeological objects, hierarchical structures found in archives (like Encoded Archival Description), and collections of objects with some relation to each other.

There are different classes in the Europeana Data Model where hierarchies can be represented:



**Figure 9.** Digital object schema in the Europeana Data Model using an event-based approach.

‘edm:ProvidedCHO’, ‘edm:WebResource’ and context classes (represented by ‘edm:Agent’, ‘edm:Place’, ‘edm:TimeSpan’, ‘edm:Concept’ or ‘edm:Event’). Vertical relationships can be expressed with the properties ‘dcterms: isPartOf’ and ‘dcterms: hasPart’. For horizontal relationships, there is the property ‘edm:isNextInSequence’, which allows for the ordering of parts. Other relationships can be explicitly expressed by adding the property ‘edm:isRelatedTo’ or its subproperty ‘edm:isSimilarTo’, narrowed by the property ‘edm:isDerivativeOf’.

Regarding the aggregation of hierarchical objects in Europeana, the providing institution must take into account several issues, such as the granularity of the hierarchy, the detail of the description and the propagation of metadata between different levels (inheritance), as well as the search, presentation and navigation through the different levels of the objects (Bardi et al., 2014).

**Multilingualism and metadata enrichment.** According to Stiller et al. (2013), there are four levels of multilingualism: display; search and navigation; presentation and translation of results; and user engagement. At Europeana, the static content of its pages is available in 31 different languages, which can be selected by the user at first access and automatically kept on subsequent visits. The availability of search terms in several languages stems from the detail of the metadata provided by each collaborating institution. The discovery of related resources is available mainly from Europeana metadata enrichment (subject, date, place, author) with multilingual vocabularies. This process involves normalizing the data provided and automatically associating links with controlled terms

in its data set or Linked Data vocabularies from the textual information provided. This process is called ‘semantic enrichment’ and corresponds to the creation of context classes according to the Europeana Data Model.

The enrichment of subject metadata, with all the translations of a concept, allows the retrieval of documents in languages other than that used to perform the search. However, and also according to Stiller et al. (2013), qualitative studies show that semantic enrichment can introduce errors if not implemented well.

In Europeana, navigation is key to enabling its users to move through its multiple collections and explore unknown content in languages often unfamiliar to them. However, language-independent navigation, available through a timeline or a map, can only be carried out in the search results. Europeana provides several facets for filtering results, including language, but sometimes it is not clear whether it corresponds to the language of the viewed object or the language in which the metadata is provided (Stiller et al., 2013).

To engage users, Europeana provides virtual exhibits on highly relevant topics in several languages. It also enables the translation of search results and objects through an automatic translation system. It creates new information for the resources it integrates, with the aim of adding value for its users (Stiller et al., 2013). Matching Europeana’s object metadata with external semantic data creates links between resources. The links created point to additional data, such as translated terms or broader terms. Europeana enriches place-related information based on GeoNames, while people’s names

and concepts are enriched with DBpedia and many other vocabularies.<sup>24</sup> As mentioned above, the Europeana Data Model has context classes so that concepts can be incorporated from vocabularies such as thesauruses, authority lists and classification systems, whether they originate from the institutions that provide content to Europeana or from other trusted external data sources.

In December 2021, Europeana launched its medium-term strategy to improve the availability of multilingualism on its platform, taking advantage of technological advances in this field (Neale, 2020). Its aim is to make Europeana's website navigation, textual search and reading of digital objects available in multiple languages. This will require the use of reliable vocabularies with multilingual coverage of the existing metadata and the translation of the metadata and text into English, which will act as a hinge language for areas not covered by the vocabularies. Europeana will also use professional translators or simultaneous translation services to make the user interface, the website, the metadata and the full-text content available in multiple languages (Neale, 2020).

### *Integration of resources from libraries, museums and archives*

The Europeana Data Model allows the integration of data from multiple sources received in various formats and the mapping of metadata standards – for example, the MARC 21 format used in libraries; the Metadata Encoding and Transmission Standard for digital libraries;<sup>25</sup> Encoded Archival Description, which is used in archives; and Lightweight Information Describing Objects, which is used in museums. To integrate this content into Europeana, these formats are mapped and converted to the Europeana Data Model using RDF/XML serialization. The data provided is normalized and semantically enriched through links to specific vocabularies.

To share digital collections with Europeana, the scope of the institution's collections must be aligned with the Europeana Content Strategy (Scholz et al., 2017) and meet the technical criteria defined in the Europeana Publishing Guide.<sup>26</sup> To publish digital content, it is necessary to find an aggregator, this being an organization that gathers data and makes it accessible through Europeana.

At the University of Coimbra, we have a vast collection of digital cultural heritage content available in our digital library, Almamater, which is shared with Europeana through our national aggregator, the Registo Nacional de Objetos Digitais, managed by the National Library of Portugal.<sup>27</sup> Briefly, for each

digital object, we register links to the corresponding item in the library catalogue, which will be converted into the 'edm:ShownBy' and 'edm:ShownAt' properties of the Europeana Data Model. The metadata from our library catalogue is converted from MARC 21 to UNIMARC and imported into the Registo Nacional de Objetos Digitais platform using the ISO 2709:2008 exchange format (International Organization for Standardization, 2008). The data is checked and sent to Europeana, processed and pre-published for validation. After validation, the data is finally published. Depending on the quality of the metadata and images, there are four tiers of criteria related to the quality of the content and three tiers to the quality of the metadata (for more information, see Scholz, 2015).

The University of Coimbra has approximately 1200 digital objects in Europeana, which link to the International Image Interoperability Framework viewer in Almamater to display the images and navigate the document structure.<sup>28</sup> This is due to the fact that the International Image Interoperability Framework resource mapping is not yet available on the Registo Nacional de Objetos Digitais platform, but we hope to embed this presentation mode in our Europeana objects soon.

### *Europeana developments*

The Europeana Data Model is a flexible and interoperable model for the description of cultural heritage resources. It is the result of systematic collaborative work and has the possibility to be expanded. This collaborative work has made it possible to accommodate the particularities of resources in specific areas through the development of extensions and adjustments that can respond to these characteristics. Europeana facilitates this process by supporting the community through the provision of detailed documentation on the Europeana Data Model, the developments being made, and the definition of the Europeana Data Model profiles (changes to the base model) to support future data integration services on the platform.

Some of the profiles that have been defined are the Sound Profile, to support the characteristics of sound objects and the metadata used to describe them; the International Image Interoperability Framework for the Europeana Data Model to provide guidelines for institutions for including resource descriptions in the International Image Interoperability Framework format in the Europeana Data Model metadata; and the Annotations Profile, based on the W3C Web Annotation Data Model recommendation (World Wide Web Consortium, 2017), which provides a solution for

representing annotations so that they can be shared and used by different platforms related to Europeana. Storytelling is also an area in which Europeana is investing by developing specific methods and techniques together with its partners.

Europeana also offers a wide range of application programming interfaces (APIs),<sup>29</sup> allowing the creation of applications using the rich cultural heritage available in Europeana: the Search API for simple searches; the SPARQL service to search Europeana's structured metadata; the Record API to retrieve the metadata of a specific record; the Entity API to access information about subjects, people and places; the Annotations API to add information about existing items in Europeana; and, finally, there is also the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) service to retrieve larger volumes of information or aggregate Europeana's repository (Europeana Foundation, 2008).

## Conclusions and future work

Linked Data models have been developed and adopted in various areas, and their application is crucial for cultural heritage institutions. The adoption of Linked Data models makes it possible to disambiguate searches and link all kinds of relevant information, which can be updated dynamically, allowing for the creation of a wide network of resources. When data is linked in this way, it is easier to discover and share information, which in turn is more complete, and allows one to create applications based on one's own and other people's data. Recognizing the importance of Linked Data models, several standards organizations in the field of libraries, museums and archives have developed conceptual models and information representation formats that follow these principles. Europeana is one of the platforms that uses a Linked Data model – the Europeana Data Model. It aggregates digital cultural heritage content from various types of institutions and allows the representation of all kinds of digital objects associated with cultural heritage. Europeana includes many millions of digital objects and has developed a search portal and a set of services that enable its repository and its users to grow. The semantic enrichment of metadata, the creation of thematic collections and the provision of a set of services to the institutions participating in Europeana, in addition to an active community of partners, has contributed to the success of this platform.

However, there are some challenges in this area. The volume of the digitization of objects in cultural heritage institutions is increasing. There is growing concern about the creation of detailed metadata,

high-quality digitization and reproduction rights to enable the use of assets in the cultural heritage context and in other sectors. Enhancing the educational value of digital cultural heritage through the development of specific programmes and content should be considered. The collaborative reuse of digital cultural heritage is also an area to be developed, with the implementation of partnership projects using shared platforms and innovative tools to reach users more quickly and provide better-quality experiences.

The undeniable advantage in adopting Linked Data models poses numerous challenges for cultural heritage institutions to adapt their registration systems or, at the very least, the presentation of information. In addition to taking advantage of Linked Data as a source of data created outside an institution, it can also make an invaluable contribution by exposing an institution's thesauruses and other authority structures for use by other communities. The lack of commercial solutions and the available open-source solutions require a huge adaptation and development effort to make its implementation possible, which necessarily involves the setting up of multidisciplinary teams to put this transformation into practice. Such challenges will also require the creation of partnerships between different institutions, allowing the establishment of joint solutions and thus making up for the lack of skilled human resources, particularly in the field of information and communications technology (Paquet, 2020). We hope that, in this way, it will be possible for the millions of digital resources of cultural heritage institutions to soon emerge from the silos in which they currently reside and become freely discoverable, available and searchable on the Semantic Web.

## Acknowledgement

The authors wish to thank Professor Maria Cristina Freitas in the Faculty of Arts and Humanities at the University of Coimbra for her observations and suggestions on an earlier version of this article.



## Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

## Funding

The authors received no financial support for the research, authorship and/or publication of this article.

## ORCID iDs

Ana Luísa Silva  <https://orcid.org/0000-0001-6909-0700>  
 Ana Lúcia Terra  <https://orcid.org/0000-0003-1292-2849>

## Notes

1. Throughout the article, the broader term Linked Data will be used.
2. Europeana is a virtual library that was developed on the initiative of the European Union. See <https://www.europeana.eu>
3. W3C is an international community that develops open standards for the Web. See <https://www.w3.org>
4. HTML is a set of Web technologies that includes HyperText Markup Language. Hyperlinks are references to data or documents on the Web that a user can access by selecting the link.
5. A URI is a unique identifier for a Web resource. HTTP is the basis for data communication through the Web.
6. RDF/XML (World Wide Web Consortium, 2014c); N3 (World Wide Web Consortium, 2011); Turtle – textual syntax for RDF (World Wide Web Consortium, 2014b); N-Triples – line-based syntax for RDF graphs (World Wide Web Consortium, 2014a); RDF/JSON – recommended serialization for RDF (World Wide Web Consortium, 2013).
7. For SPARQL, see World Wide Web Consortium (2008).
8. For the SKOS, see World Wide Web Consortium (2009). OWL is a language for representing knowledge about things, groups of things and relationships between things (World Wide Web Consortium, 2012).
9. foaf (friend of a friend) is a machine readable ontology describing persons, their relations with objects and other people and also their activities.
10. See Wikipedia at <https://en.wikipedia.org/>
11. The Dublin Core is a vocabulary for Web document description maintained by the Dublin Core Metadata Initiative (2020).
12. SPARQL endpoint is a web service that allows querying sets of RDF triples using the SPARQL semantic query language.
13. Digital Public Library of America (<https://dp.la/>); DBpedia (<https://www.dbpedia.org/>); Wikidata (<https://www.wikidata.org/>).
14. MARC 21 formats “are standards for the representation and communication of bibliographic and related information in machine-readable form” (Library of Congress, 2023).
15. The Metadata Authority Description Schema in RDF is a knowledge organization system for authority control, thesauruses, taxonomies, subject heading systems and others. It is a widely used RDF vocabulary and is related to the SKOS (Library of Congress, 2015).
16. For the Schema.org vocabulary, see <https://schema.org/>
17. For the Virtual International Authority File, see <https://viaf.org/>
18. Lightweight Information Describing Objects is an XML metadata format to deliver information on museums’ objects in a standardized way.
19. Registo Nacional de Objetos Digitais (<https://rnod.bnportugal.gov.pt/rnod/>); Archives Portal Europe (<https://www.archivesportaleurope.net/>); MUSEUHUB (<http://www.museuhub.eu/>); Manuscriptorium (<http://www.manuscriptorium.com/>); Musical Instruments Museums Online (<https://mimo-international.com/MIMO/>); European Fashion Heritage Association (<https://fashionheritage.eu/>); Technische Informationsbibliothek (<https://www.tib.eu/en/>).
20. The Dublin Core™ Metadata Element Set, Version 1.1 (Dublin Core Metadata Initiative, 2012) was replaced by the DCMI metadata terms (Dublin Core Metadata Initiative, 2020).
21. For the Deutsche Digitale Bibliothek, see <https://www.deutsche-digitale-bibliothek.de/>
22. For ROSSIO, see <https://rossio.pt/front/home>
23. With the Europeana Data Model, Europeana can incorporate complex objects: archival holdings can be represented as a whole and with their individual components, such as letters, deeds or manuscripts. Also, for digitized books, their chapters, indexes or illustrations can be represented as a whole or individually.
24. GeoNames is a geographical database with every country that contains millions of place names (see <http://geonames.org>). DBpedia is the main Linked Open Data resource on the Web and contains over 228 million entities (see <http://dbpedia.org>).
25. For the Metadata Encoding and Transmission Standard, see <https://www.loc.gov/standards/mets/>
26. For the Europeana Publishing Guide, see <https://europeana.atlassian.net/wiki/spaces/EF/pages/2059763713/EPF+-+Publishing+guidelines>
27. For Almamater, see <https://am.uc.pt>
28. The International Image Interoperability Framework is a framework for delivering image-based content on the Web. See <https://iiif.io/>
29. An application programming interface is software that has been created according to a set of rules that enables machine-to-machine communication.

## References

- Agenjo K, Arnold X, Bailly R, et al. (2021) Recommendations for the representation of hierarchical objects in Europeana. Europeana pro, 12 August. Available at: <https://pro.europeana.eu/project/hierarchical-objects> (accessed 7 November 2022).
- Archer P and Perego A (2012) Core Person Vocabulary. Available at: <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/core-person-vocabulary/distribution/person-v100rdf> (accessed 7 February 2022).
- Bekiari C, Doerr M, Le Bœuf P, et al. (eds) (2015) FRBR: Object-oriented definition and mapping from FRBR<sub>ER</sub>, FRAD and FRSAD (version 2.2). Available at: [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr/frbroo\\_v2.2.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr/frbroo_v2.2.pdf) (accessed 12 February 2023).
- Berners-Lee T (2006) Linked Data. Available at: <https://www.w3.org/DesignIssues/LinkedData.html> (accessed 4 February 2022).
- Ciocoiu A (2018) Introducing the Europeana Data Model online course on data aggregation. Europeana pro, 18

- January. Available at: <https://pro.europeana.eu/post/introducing-the-europeana-data-model-online-course-on-data-aggregation> (accessed 7 February 2022).
- Dublin Core Metadata Initiative (2012) Dublin Core™ Metadata Element Set, Version 1.1: Reference description. Available at: <https://www.dublincore.org/specifications/dublin-core/dces/>
- Dublin Core Metadata Initiative (2020) DCMI metadata terms. Available at: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- Europeana Foundation (2008) Services and tools. Available at: <https://pro.europeana.eu/about-us/services-and-tools> (accessed 28 December 2021).
- Europeana Foundation (2014) Europeana Data Model. Available at: <https://pro.europeana.eu/page/edm-documentation> (accessed 7 February 2022).
- Expert Group on Archival Description (2021) Records in Contexts—Conceptual Model. International Council on Archives, 22 July. Available at: <https://www.ica.org/en/records-in-contexts-conceptual-model> (accessed 7 February 2022).
- Herman I (2010) Tutorial on Semantic Web technologies. Available at: <https://www.w3.org/People/Ivan/CorePresentations/RDFTutorial/Slides.html> (accessed 25 January 2022).
- International Council of Museums (n.d.) CIDOC Conceptual Reference Model. Available at: <https://www.cidoc-crm.org/> (accessed 7 February 2022).
- International Organization for Standardization (2008) *Information and documentation – Format for information exchange*. ISO 25964-2:2013. Geneva: International Organization for Standardization.
- International Organization for Standardization (2023) *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*. ISO 25964-2:2013. Geneva: International Organization for Standardization.
- International Organization for Standardization (2020) *Information and documentation – A reference ontology for the interchange of cultural heritage information*. ISO 21127:2014. Geneva: International Organization for Standardization.
- Isaac A and Clayphan R (2013) Europeana Data Model primer. Available at: [https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Primer\\_130714.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf)
- Joudrey DN and Taylor AG (2018) *The Organization of Information*. 4th ed. Santa Barbara, CA: Libraries Unlimited.
- Library of Congress (n.d.) ID.LOC.GOV – Linked Data service. Available at: <https://id.loc.gov/>
- Library of Congress (2015) MADS/RDF primer. Available at: <https://www.loc.gov/standards/mads/rdf/> (accessed 3 February 2022).
- Library of Congress (2023) Marc Standards. Available at: <https://www.loc.gov/marc/>
- Neale A (2020) Europeana DSI-4 multilingual strategy. Europeana Pro, 12 October. Available at: <https://pro.europeana.eu/post/europeana-dsi-4-multilingual-strategy> (accessed 28 December 2021).
- Paquet AP (2020) Linked data and linked open data projects for libraries, archives and museums: Constructing pathways to information discovery and cultural heritage sector collaboration. Available at: <https://jscholarship.library.jhu.edu/handle/1774.2/63875>
- Park H and Kipp M (2019) Library Linked Data models: Library data in the Semantic Web. *Cataloging and Classification Quarterly* 57(5): 261–277.
- Petras V, Hill T, Stiller J, et al. (2017) Europeana – A search engine for digitised cultural heritage material. *Datenbank-Spektrum* 17(1): 41–46.
- Raafat T (2017) RDF and OWL: The powerful duo. Available at: <https://connecteddataworld.com/rdf-and-owl-the-powerful-duo>
- Scholz H (2015) Publishing Framework. Europeana Pro, 1 September. Available at: <https://pro.europeana.eu/post/publishing-framework>
- Scholz H, Devarenne C, Freire N, et al. (2017) Europeana Content Strategy. Available at: <https://pro.europeana.eu/post/europeana-content-strategy>
- Silva GM, Glória AC, Salgueiro AS, et al. (2022) ROSSIO infrastructure: A digital humanities platform to explore the Portuguese cultural heritage. *Information* 13(2): Article 50.
- Stein R and Balandi O (2019) Using LIDO for evolving object documentation into CIDOC CRM. *Heritage* 2(1): 1023–1031.
- Stiller J, Gäde M and Petras V (2013) Multilingual access to digital libraries: The Europeana use case. *Information – Wissenschaft & Praxis* 64(2–3): 86–95.
- Wang Y and Yang S (2018) Linked data technologies and what libraries have accomplished so far. *International Journal of Librarianship* 3(1): 3–20.
- Wood D, Zaidman M, Ruth L, et al. (2014) *Linked Data: Structured Data on the Web*. Shelter Island, NY: Manning.
- World Wide Web Consortium (2008) SPARQL query language for RDF. Available at: <https://www.w3.org/TR/rdf-sparql-query/>
- World Wide Web Consortium (2009) SKOS Simple Knowledge Organization System. Available at: <https://www.w3.org/TR/skos-reference/>
- World Wide Web Consortium (2011) Notation3 (N3): A readable RDF syntax. Available at: <https://www.w3.org/TeamSubmission/n3/>
- World Wide Web Consortium (2012) OWL 2 Web Ontology Language: Document overview (second edition). Available at: <https://www.w3.org/TR/owl2-overview/>
- World Wide Web Consortium (2013) RDF 1.1 JSON alternate serialization (RDF/JSON). Available at: <https://www.w3.org/TR/rdf-json/>
- World Wide Web Consortium (2014a) RDF 1.1 N-Triples: A line-based syntax for an RDF graph. Available at: <https://www.w3.org/TR/n-triples/>
- World Wide Web Consortium (2014b) RDF 1.1 Turtle: Terse RDF triple language. Available at: <https://www.w3.org/TR/turtle>

- World Wide Web Consortium (2014c) RDF 1.1 XML syntax. Available at: <https://www.w3.org/TR/rdf-syntax-grammar>
- World Wide Web Consortium (2015a) Web standards. Available at: <https://www.w3.org/standards/semantic-web/> (accessed 4 February 2022).
- World Wide Web Consortium (2015b) Rich Structured Data Markup for Web Documents. Available at: <https://www.w3.org/TR/rdfa-primer/> (accessed 4 February 2022).
- World Wide Web Consortium (2020) A JSON-based Serialization for Linked Data. Available at: <https://www.w3.org/TR/json-ld11/> (accessed 4 February 2022).
- World Wide Web Consortium (2017) Web Annotation Data Model. Available at: <https://www.w3.org/TR/annotation-model/>
- Zapounidou S, Sfakakis M, Papatheodorou C, et al. (2017) Representing and integrating bibliographic information into the Semantic Web: A comparison of four conceptual models. *Journal of Information Science* 43(4): 525–553.

### Author biographies

**Ana Luísa Silva** is a senior information technology specialist at the University of Coimbra's general library, has a degree in Computer Engineering and is responsible for the technical management and coordination of the university

libraries' information systems. She has been working in the library domain since 1993. She has collaborated in several projects and working groups at the national level. Currently, she is a PhD student in Information Science in the Faculty of Arts and Humanities at the University of Coimbra and a collaborating investigator in the Centre for Interdisciplinary Studies (CEIS20) at the University of Coimbra.

**Ana Lúcia Terra** is an assistant professor in the Faculty of Arts and Humanities at the University of Coimbra. She holds a PhD in Information Science from the University of Coimbra (2009). She is also a researcher in the Centre for Interdisciplinary Studies (CEIS20). She is president of the Iberian Chapter of the International Society for Knowledge Organization and an elected member of the board of the Bobcatsss Association (formerly EUCLID – European Association for Library and Information Education and Research). Her research and publications focus on information management, information behaviour, knowledge organization and information representation. She participates as a principal investigator and researcher in several funded projects (including Erasmus+ and FCT – Fundação para a Ciência e a Tecnologia).