Omeka S as an alternative for the development of digital collections and digital humanities projects

Rubén Alcaraz-Martínez[1]

[1] *Departament de Biblioteconomia, Documentació i Comunicació Audiovisual, Universitat de Barcelona, Spain*

Abstract

Objectives: a detailed description of the main features and functionalities of Omeka S, a free software content management system aimed at creating repositories and digital libraries, as well as being widely used in various kinds of digital humanities projects, is presented.

Methodology: the analysis of the characteristics of Omeka S has been carried out from a literature review, the information available on the website and the official documentation of the application, as well as through its installation, use, and analysis. The description of the application is complemented by a compilation of various international use cases in the field of digital humanities that showcase the versatility of the tool. Finally, a comparison is offered with five other free software applications available on the market: Omeka Classic, CollectiveAccess, Mukurtu, ResearchSpace and Arches.

Results: among the main strengths of Omeka S is its ease of installation, maintenance, and use, its multisite architecture, the ease of integrating vocabularies from the semantic web as ontologies, its flexibility to create description templates for digital objects, the freedom in the design of the pages that make up the public interface, or its native integration with other repository management systems such as DSpace or Fedora Commons, thus being able to be implemented as a platform for the management and publication of resources derived from the repository, or as a public access layer of the digital objects, respectively. On the other hand, managing a project with Omeka S requires a greater initial effort due to its great flexibility. Finally, the system is less competent at accessing the database, which cannot be done natively through SPARQL queries.

Keywords

Omeka S, Digital collections, Digital repositories, Digital Humanities, Semantic Web

* ralcaraz@ub.edu

This postprint is the author's English translation of the article published in:

Alcaraz-Martínez, Rubén (2022). Omeka S como alternativa para el desarrollo de colecciones digitales y proyectos de humanidades digitales. *BiD: Textos Universitaris de Biblioteconomia i Documentació, 48.* https://doi.org/10.1344/BiD2022.48.06

1. Introduction

Omeka S (Semantic) is a free and open-source content management system designed to enable a wide range of cultural institutions—libraries, archives, museums, universities, among others—to develop digital projects of various kinds. Omeka S is a project of the Corporation for Digital Scholarship at the Roy Rosenzweig Center for History and New Media at George Mason University, an organisation also responsible for leading initiatives in their respective fields, such as Omeka Classic and the bibliographic reference management system Zotero, among others.

Omeka S forms part of a new generation of technological platforms which, unlike more "traditional" repository management systems such as ContentDM, Greenstone, DSpace, EPrints or Invenio, incorporates a wide range of functionalities, resources, and technologies for creating derivative products based on repository content. These include virtual exhibitions, timelines, and interactive resources that enable users to explore or interact with repository objects, or to access them through different modes of information visualisation (Marsh, 2017).

The limitations on publishing and generating derivative outputs that characterise the aforementioned software applications have also led to the adoption of Omeka S as a public-facing communication layer built on top of repository management systems through their respective APIs (Application Programming Interfaces). It has likewise been used as a content management system for structuring and presenting large datasets originating from third-party systems. Representative examples can be found in the works of Hoyle (2017), Laudito (2020), and Wittmann *et al.* (2019).

When examining this application, the first point to emphasise is the distinction between this software and the platform known as Omeka Classic (Alcaraz, 2012). Despite sharing a name, they are two entirely independent applications that do not share code or a development roadmap. Beyond their titles, the similarities between the two are limited to their overarching purpose: the dissemination of digital collections on the web and the publication of virtual exhibits or other types of projects within the field of digital humanities.

The aim of this study is to examine in detail the characteristics, functionalities, and capabilities of Omeka S, as well as the themes and features of the projects developed with it. Interest in this application is justified by the recognised prestige of its developers, who are responsible for software adopted by thousands of projects and institutions worldwide, and by the significant and active user community that supports it. This article focuses on version 3.0.2 of Omeka S, the latest stable release available at the time of analysis.

The article is structured as follows: the Methodology section outlines the approach used to analyse the application under study. This is followed by a detailed discussion of the architecture and technical requirements of Omeka S, the creation and configuration of a site, the management of metadata vocabularies and cataloguing templates, as well as the purpose and value of its main modules and functionalities. The Projects section presents a selection of initiatives implemented with Omeka S. Finally, a comparison is drawn between Omeka S and other applications in the field of digital repositories and digital humanities projects.

2. Theoretical framework

The methodology consisted of a bibliographic review of studies published on Omeka S and on projects implemented using this application, alongside an examination of the information available in the official Omeka S documentation, and the installation, use, and analysis of the software itself. Subsequently, other applications were selected in order to carry out a comparison between their functionalities and those of Omeka S.

The bibliographic review was conducted in Google Scholar on 10 June 2021. The search strategy was based on the following combination of terms: "Omeka S" OR "Omeka Semantic" OR "Omeka-S" OR "Omeka-Semantic". Publish or Perish was used to manage and analyse the search results.

The eligibility criteria considered studies whose primary objective was the analysis of the characteristics of Omeka S or of a project developed using this application, published in English, French,[1] Spanish, or Catalan from 1 January 2017 onwards (the release date of the first beta version of Omeka S). Based on the non-eligibility criteria, all works that merely mentioned Omeka S without analysing it or discussing a related project were excluded, as were documents of a non-scientific nature.

Figure 1 presents a summary of the bibliographic review process.

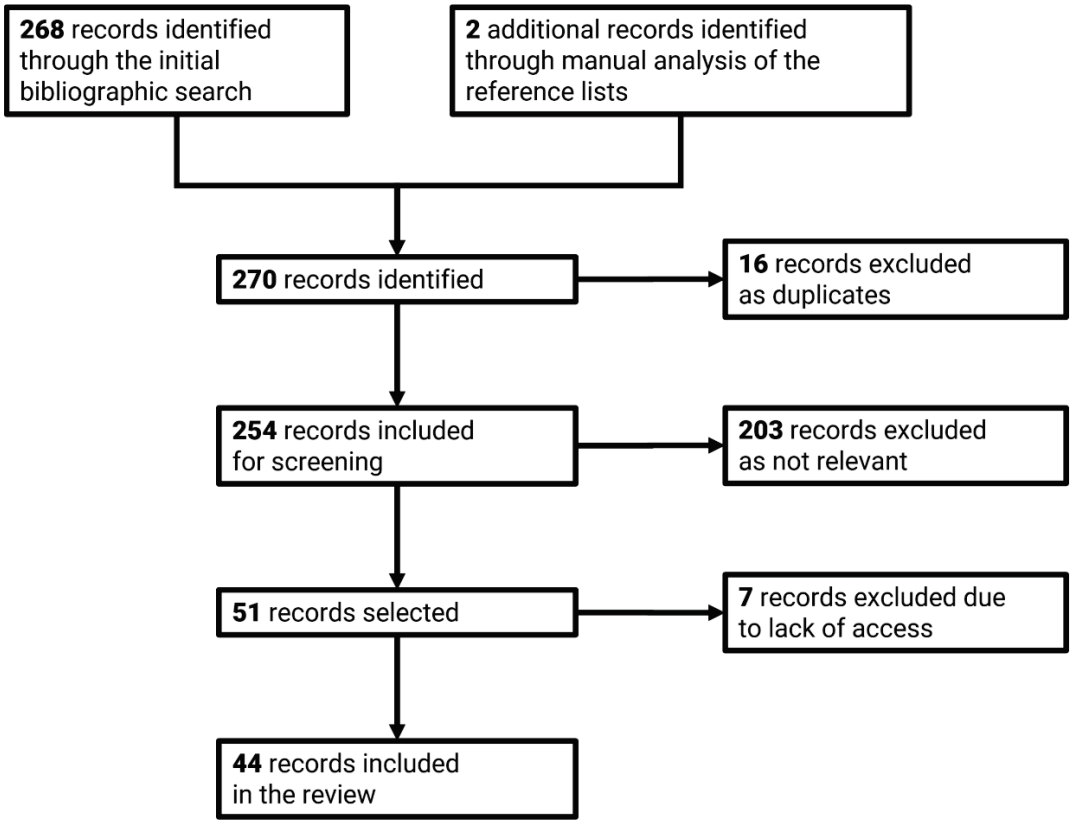

Figure 1. Diagram of the methodology used for the bibliographic review and summary of the process. Source: author's own work.

For the comparative analysis of Omeka S and other technological solutions in the same field, an analytical framework was employed based on the requirements set out in the studies by Cascón-Katchadourian *et al.* (2018) and Kubik and Kwiecień (2020). These requirements were consolidated, synthesised, and adapted to the specific features of the software examined in this study.

3. Analysis of Omeka S

3.1. Application Structure and User Management

The Omeka S administration panel is divided into two main sections (Figure 2): a block on the left-hand side of the interface containing the search tool and navigation options, and a central area where the content and specific options of the section currently in use are displayed.

[1] Both Omeka Classic and Omeka S are highly popular and widely adopted in Francophone contexts. Evidence of this is the existence of the *Association des Usagers Francophones d'Omeka*. https://omeka.fr/

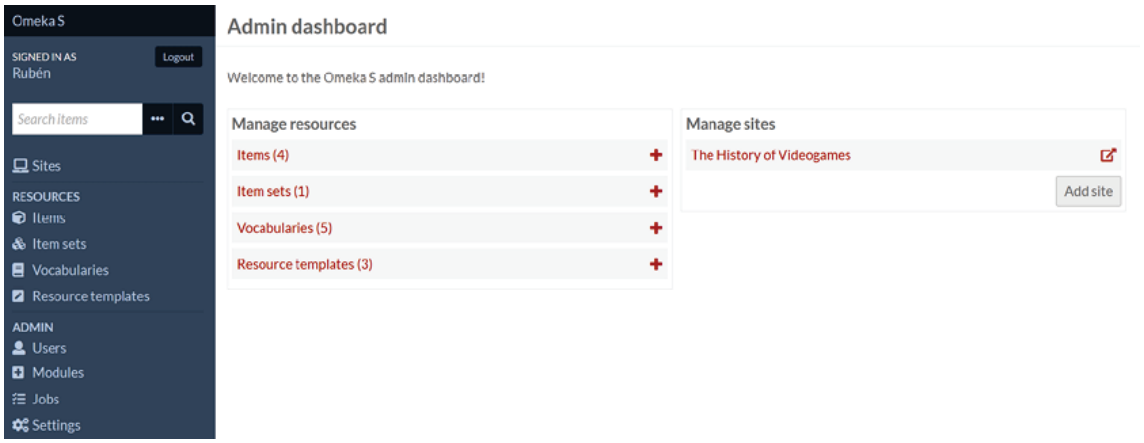

Figure 2. Home screen of the administration panel. Source: author's screenshot

The navigation menu is itself divided into three main sections: Sites, Resources, and Admin. The Sites section allows users to create and manage the various sites through which the repository's content will be publicly displayed. Under Resources, we find the options designed for creating and editing digital objects (items), the item sets that function as collections or thematic groupings of digital objects, the tools for managing the vocabularies used to describe the repository's digital objects, and the cataloguing templates. Finally, the Admin section contains user management, modules, active processes, and site settings.

3.2. Sites

Sites constitute the public-facing layer of an Omeka S instance (Omeka, 2021a). A single installation of Omeka S can host as many sites as required. Each site may correspond to a specific project and may use all or a selected subset of the objects available in the system. Every site has its own URL and features its own pages, navigation structures, user permissions, and visual appearance.

This feature makes Omeka S an attractive solution for institutions that wish to manage multiple distinct sites while sharing objects across them (Nishanbaev, Champion & McMeekin, 2019; Boyles, 2020; Li, 2020; Popović, Škorić & Rujević, 2021). Its multisite architecture is particularly useful for GLAM institutions in which different types of information units coexist under the same organisational umbrella, enabling the creation of dedicated environments for each of them—vocabularies, cataloguing templates, information-visualisation templates, differentiated public interfaces—while sharing a single administration panel through which all digital objects can be accessed.

This integration can also be leveraged in the public interface, where all kinds of relationships between objects may be established and global searches across all of them may be enabled. However, if the aim is to provide fully separate management environments for different projects within the same institution, additional Omeka S instances must be installed.

Through the options available under the Pages section of each site, it is possible to create the sections and pages that will make up the public interface. Page editing is based on a system of content blocks that offer a relatively flexible structure, allowing users to choose from various options: titles, HTML code blocks, embedded content, the display of an item together with its metadata, a summary block, and others (Figure 3). To configure a page, one simply selects the desired block and defines its content—either created on the spot or retrieved from the repository.

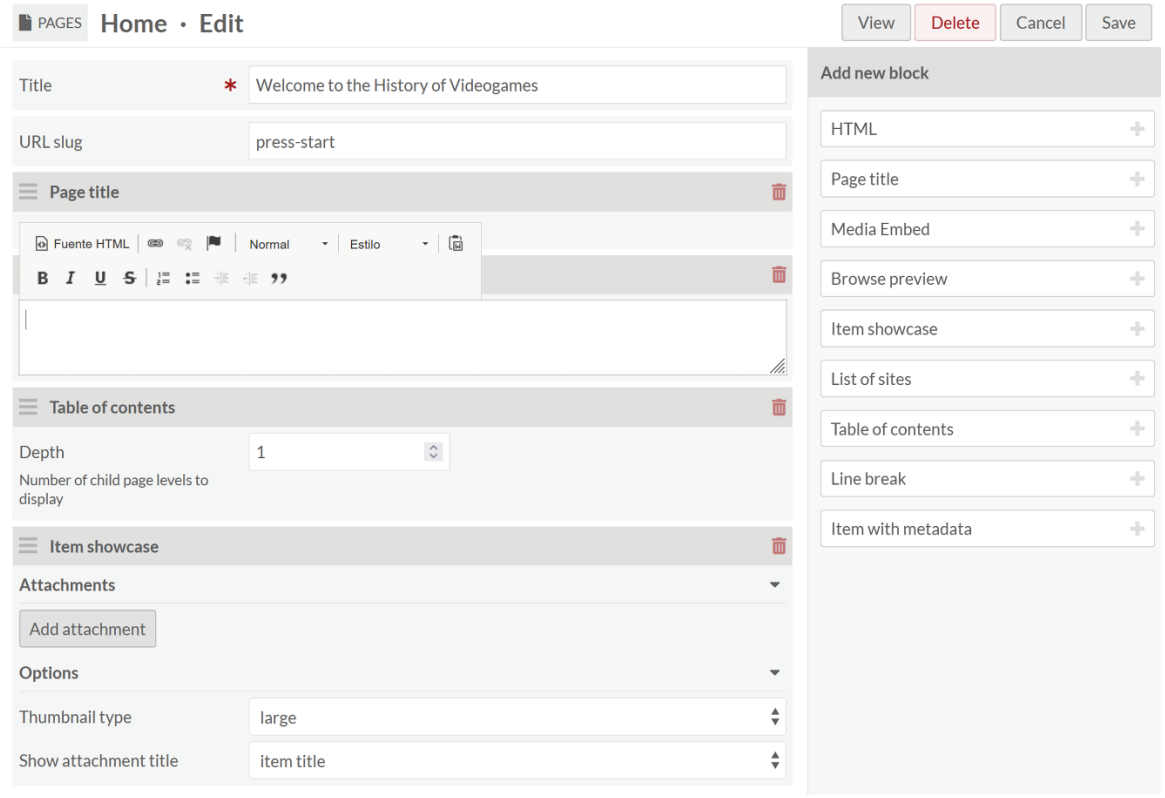

Figure 3. Page editing form. The content blocks available are shown on the right-hand side of the interface, and the editing options are displayed in the central area. Source: author's screenshot.

With regard to bibliographic records and search results pages, the application generates these on the fly, meaning that there is no need to create them manually unless one wishes to display multiple digital objects or other specific content on a particular page.

3.3. Vocabularies and Templates for Describing Digital Objects

One of the key features of Omeka S is its use of vocabularies from the semantic web for describing repository objects (Omeka, 2021b). Unlike other platforms, Omeka S does not rely on XML-based vocabularies; instead, it uses ontologies to describe repository objects. Whereas XML specifications employ a hierarchical model to represent relationships between objects and their properties, Omeka S uses the graph-based model of RDF (Resource Description Framework), in line with the principles of linked data (Popović, Škorić & Rujević, 2021).

The core application includes the following preloaded vocabularies: Dublin Core, Dublin Core Type, Bibliographic Ontology, and Friend of a Friend. These can be extended by importing any vocabulary serialised in one of the following formats: JSON-LD, N-Triples, Notation3, RDF/XML, or Turtle (Figure 4).

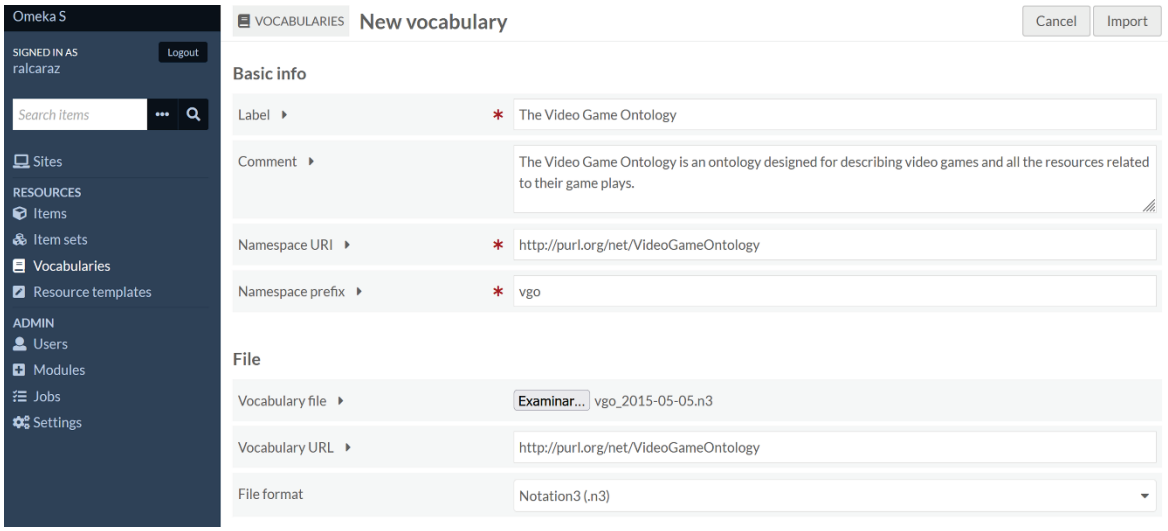

Figure 4. Form for importing a new vocabulary. Source: author's screenshot

Cataloguing templates (resource templates) are sets of classes and properties drawn from the various vocabularies available in the system, designed to facilitate the description of repository objects. By default, the system includes a base template that incorporates all the elements required by the Digital Public Library of America. Beyond this predefined template, users may create additional templates for describing different types of objects or for specific sites or projects. Each template may include as many classes and attributes from as many vocabularies as necessary (Figure 5).

It is also possible to define various aspects of each element, such as whether it is required, whether it should be public (the default setting) or private, and its data type (literal, URI, or another object available in the repository) (Figure 6).

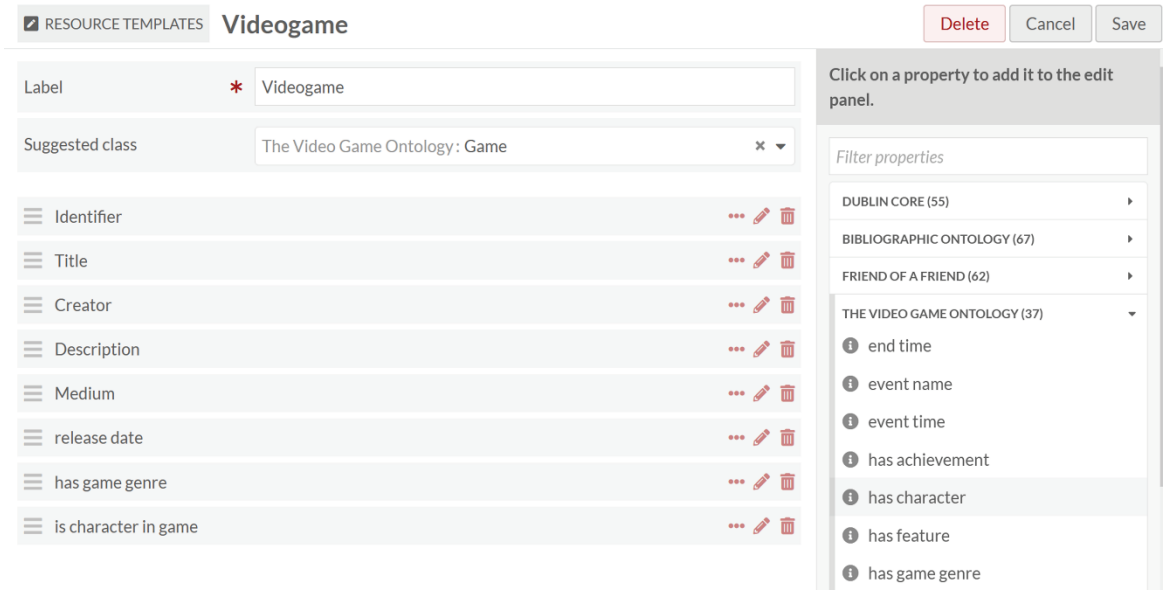

Figure 5. Form for creating a cataloguing template. The example shows a template with attributes from the Dublin Core and Video Game Ontology vocabularies. Source: author's screenshot

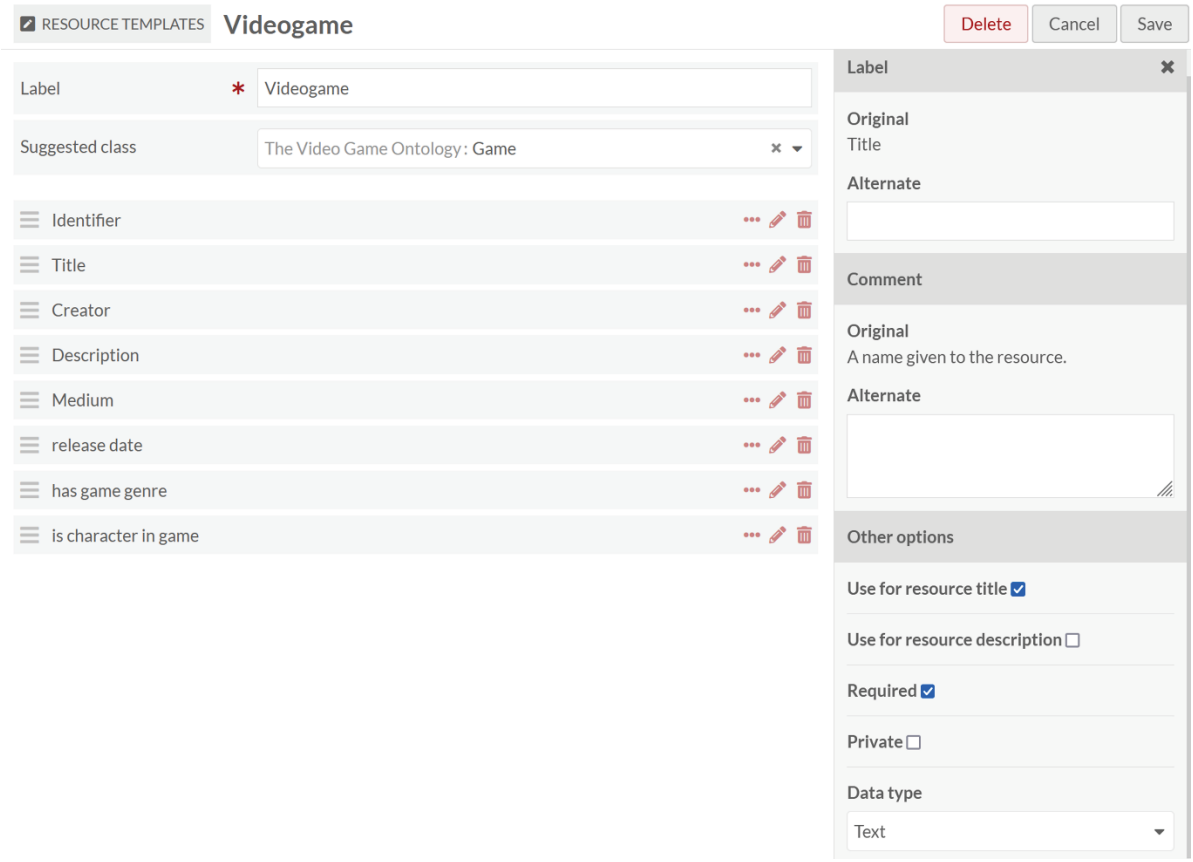

Figure 6. For each attribute, it is possible to define the public label that will appear in the object's record, as well as indicate whether the attribute represents the title or description, whether it is mandatory, and the type of data it stores. Source: author

Once a new template has been created, it becomes available for describing objects added to the repository. Working with templates does not restrict the ability to add additional attributes when necessary for a particular digital object. As shown in Figures 7 and 8, attribute values may consist of text literals, other repository objects, or a URI.

After an object has been described, it may be associated with any number of files or resources (images, PDF documents, URLs, custom HTML code, a YouTube video, etc.). Figure 9 provides an example of how a record is displayed.

Digital objects within the repository may be grouped into item sets, which follow the usual logic of collections. Unlike Omeka Classic, a digital object may belong to multiple collections and appear in multiple sites simultaneously.

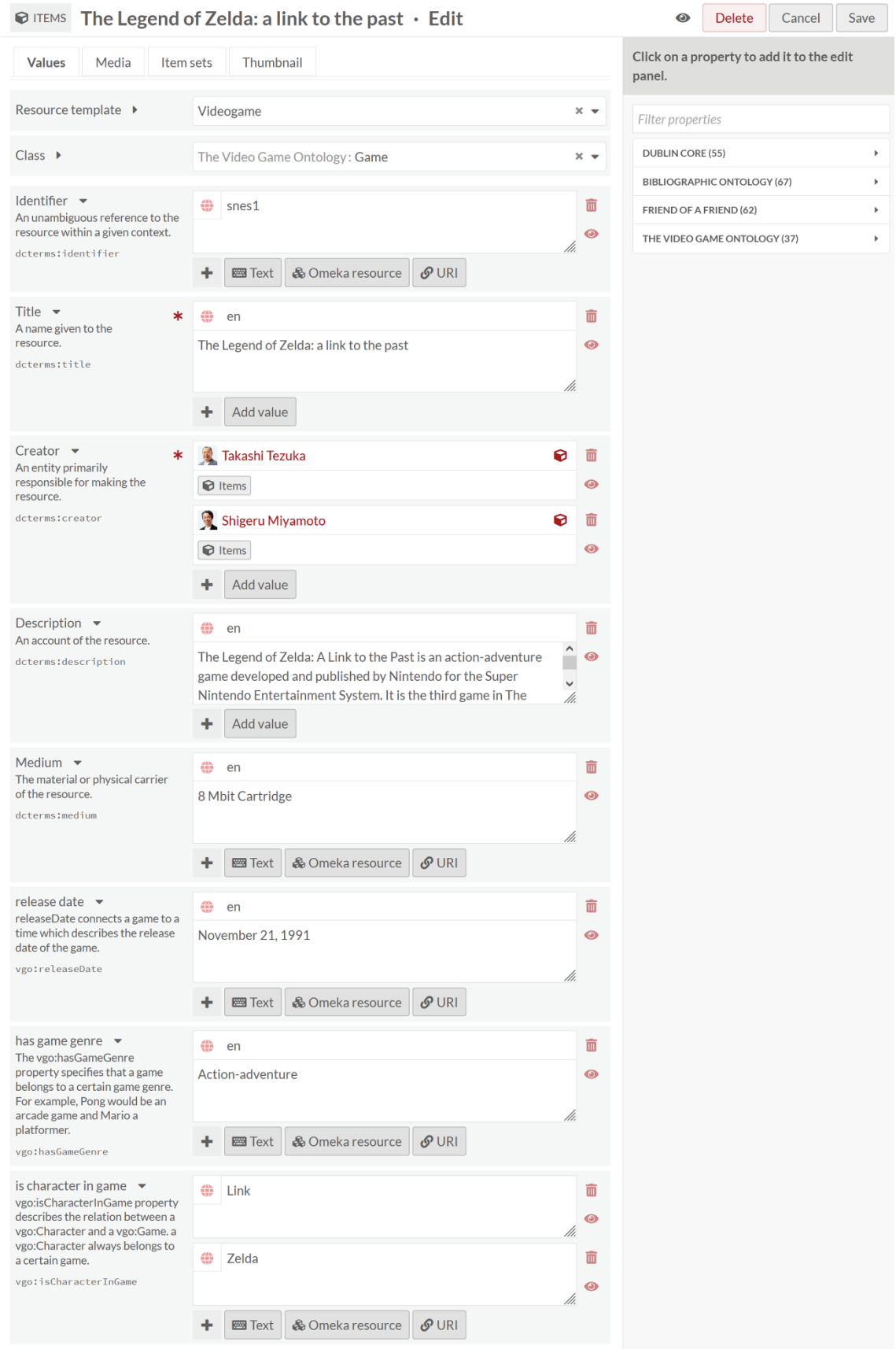

Figure 7. Form for creating a resource or item. The creator attribute has been recorded by linking this digital object to two other objects in the repository. Source: author's screenshot

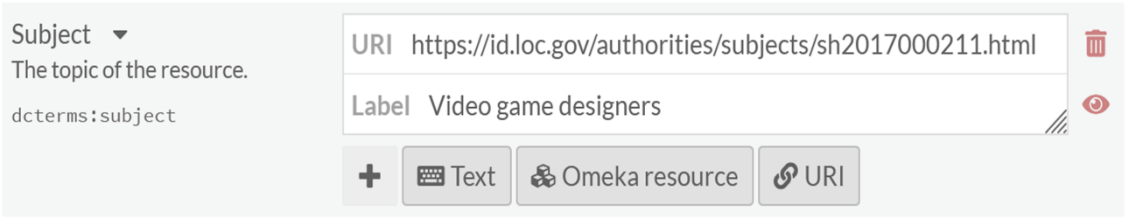

Figure 8. Subject described using LC Subject Headings and identified by its URI. Source: author's screenshot

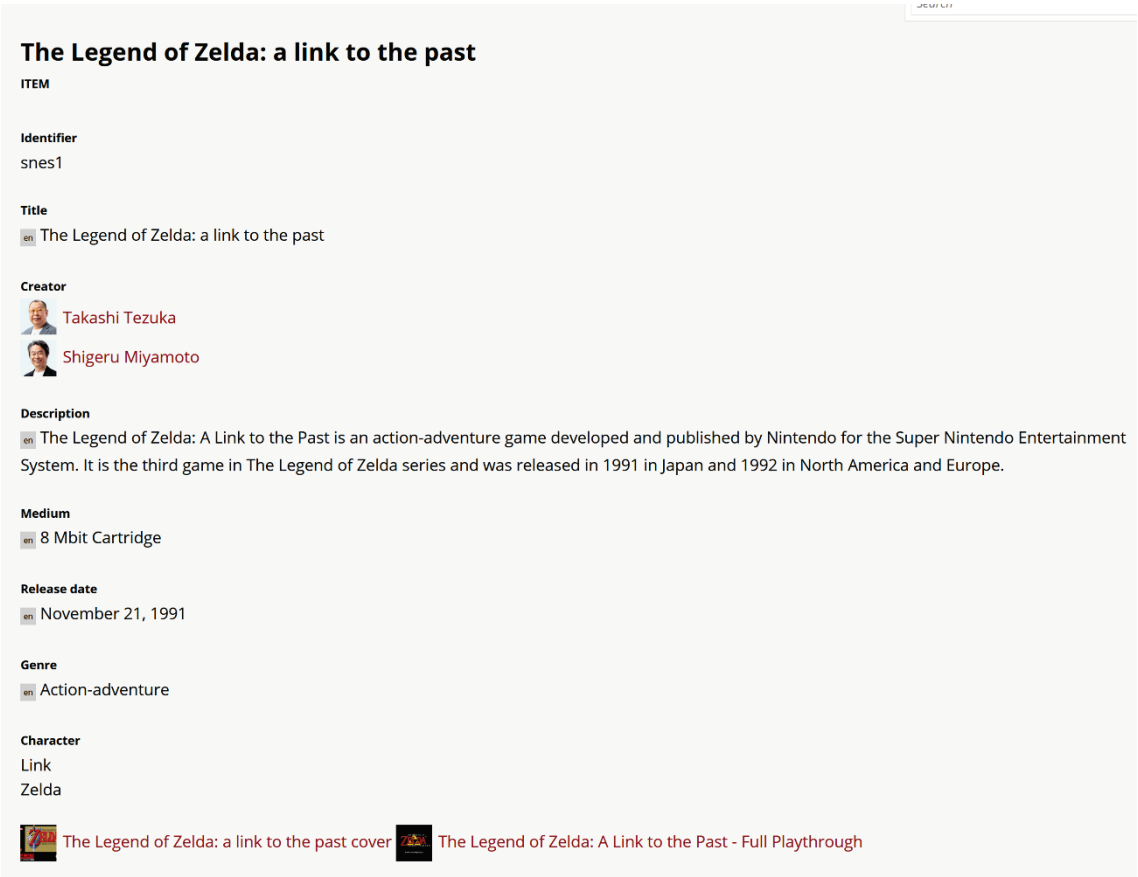

Figure 9. Example of a record display. The resources associated with the record (authors, image, and video) appear as links to other pages. Source: author's screenshot

3.4. Information Retrieval System and Interoperability

Omeka S includes a search engine that allows both simple and advanced searches, filtering by the various classes and attributes of the available vocabularies. Despite its adoption of semantic web technologies, the data are stored in a relational MySQL database, which is incompatible with direct querying through SPARQL (SPARQL Protocol and RDF Query Language) over RDF graphs (Bruneau, 2020). This represents a significant limitation for an application that positions itself within the linked data ecosystem.

One of the primary aims of Omeka S is to facilitate system integration (Leon, 2016). To this end, the application has, since its initial release, offered specific modules to support data import from systems such as DSpace or Fedora Commons through a RESTful API. In terms of interoperability, Omeka S provides a dedicated module for implementing the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). However, unlike Omeka Classic, it currently allows only the exposure of metadata, not the harvesting of metadata from third-party sources.

3.5. Modules

A defining feature of both Omeka Classic and Omeka S is that their core installations include only a small set of essential functionalities required by any digital library. Nevertheless, Omeka S offers a catalogue of modules that can be added to extend its basic capabilities. At present, the official website lists fifty-seven modules, providing functionalities such as controlled vocabulary management, working with georeferenced data, enhancing the search engine, improving resource visualisation, and enabling crowdsourcing processes through collaborative transcription or by allowing users to contribute various types of content to the repository.

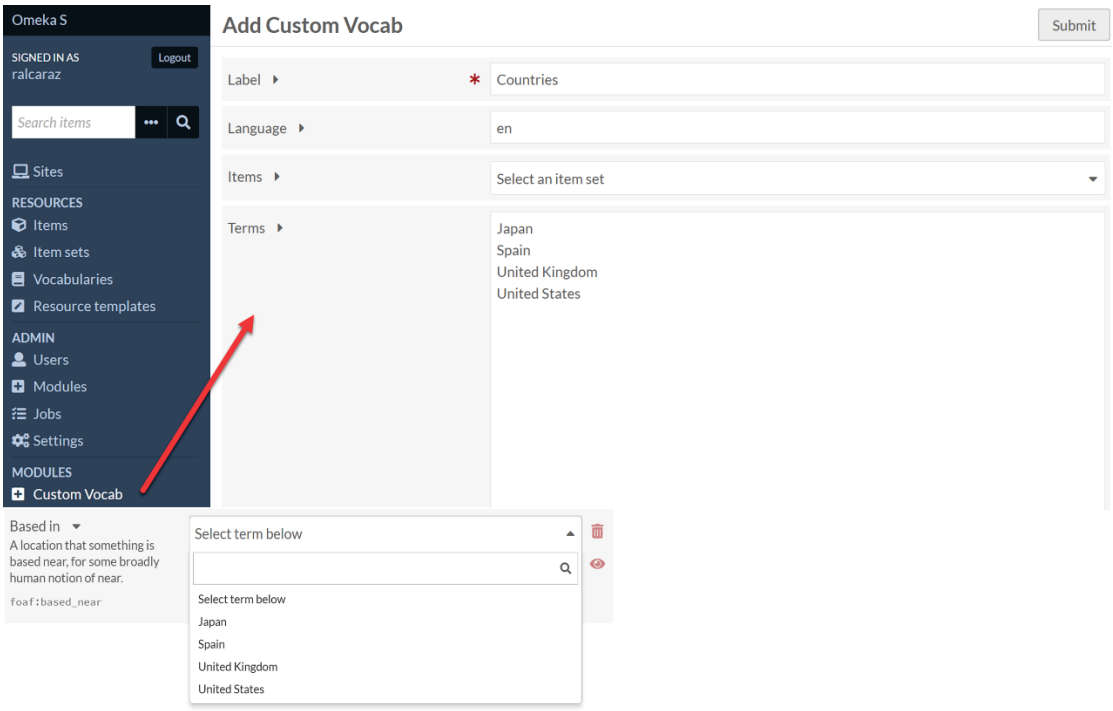

Figure 10. Creation of a custom controlled vocabulary using the Custom Vocab module and its association with a property from the FOAF ontology. Source: author's screenshot

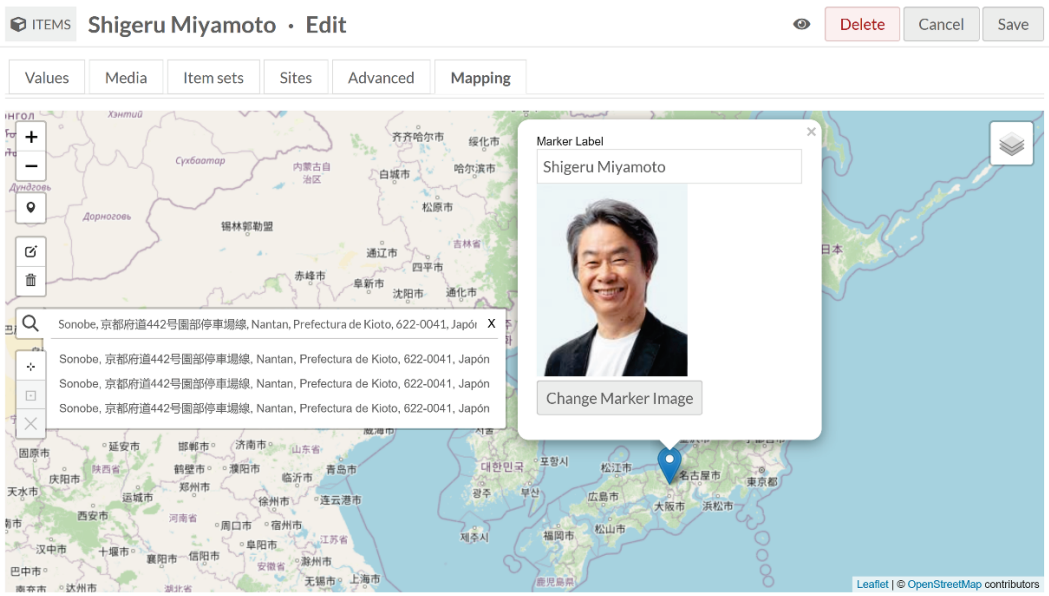

Figure 11. Interface of the module used to add geographic information to digital objects. Source: author's screenshot

3.6. Projects

Whereas Omeka Classic is commonly used for both digital humanities projects and library or archival repositories, the adoption of Omeka S is particularly prominent in the former. Representative examples include: the iconographic encyclopaedia of Christianity in the Middle Ages developed by the Institut National d'Histoire de l'Art (INHA) (Biay, Courtin & Marchesin, 2018); the correspondence archive of the French mathematician Henri Poincaré (Bruneau et al., 2020; Lasolle, Bruneau & Lieber, 2020; Lasolle *et al.*, 2021b); the cartographic visualisations enriched with semantic annotations produced by the Oronce Fine project (Benoît et al., 2018); the work of the Centro de Estudios Históricos of El Colegio de Michoacán following the digitisation, transcription, and interpretation of a corpus of legal texts from the Spanish Empire (Gayol & Melo, 2018); and oral history projects such as the Puerto Rico Disaster Archive (Boyles, 2020), created to collect personal stories about the devastation caused by Hurricane María in 2017; the ǂKhomani San Hugh Brody Collection, which gives voice to the descendants of the Indigenous hunter-gatherers of the Kalahari Desert, speakers of several endangered languages (Jones & Muftic, 2020); or Journal of the Plague Year: An Archive of COVID-19 (JOTPY) at Indiana University (Kelly & Horan, 2020; Kelly, 2020).

Other noteworthy initiatives include the repository designed to document and disseminate the excavation process and archaeological findings from the Tobiotsuka Kofun burial mound in Japan, following the CIDOC-CRM reference model (Lombardo *et al.*, 2020a); the collection of Cartesian cosmological illustrations at the University of Bucharest (Dobre, Babeș & Bujor, 2020); the creation of a digital version of the nineteenth-century French publishers' dictionary (Geslot & Rebolledo-Dhuin, 2020); thematic sites such as Slavery Images (Lovejoy, 2020), which provides teachers, researchers, and students with resources related to the African slave trade; and the database underpinning the semantic network Otletosphère, which visually represents the relationships of the Belgian documentalist Paul Otlet with family members, collaborators, works, events, and institutions (Deuff *et al.*, 2020). Omeka S has also been used to develop name authority catalogues, such as the Western Name Authority File at the University of Utah (Neatrour & Myntti, 2019).

In the Spanish context, examples include: the virtual exhibitions of the Library of the University of Seville, initially developed with Omeka Classic (Mensaque-Urbano, 2017) and migrated in 2018 to Omeka S (Peñalver & Bocanegra, 2018); the Virtual Museum of the University of Barcelona, where Omeka S replaced the previous management system, MuseumPlus; the institutional graphic archive of the University of Las Palmas de Gran Canaria, which presents its graphic and audiovisual collections; and Los Guanchismos, a dictionary of Canary Islands toponymy that explores and relates Indigenous toponyms preserved through oral tradition.

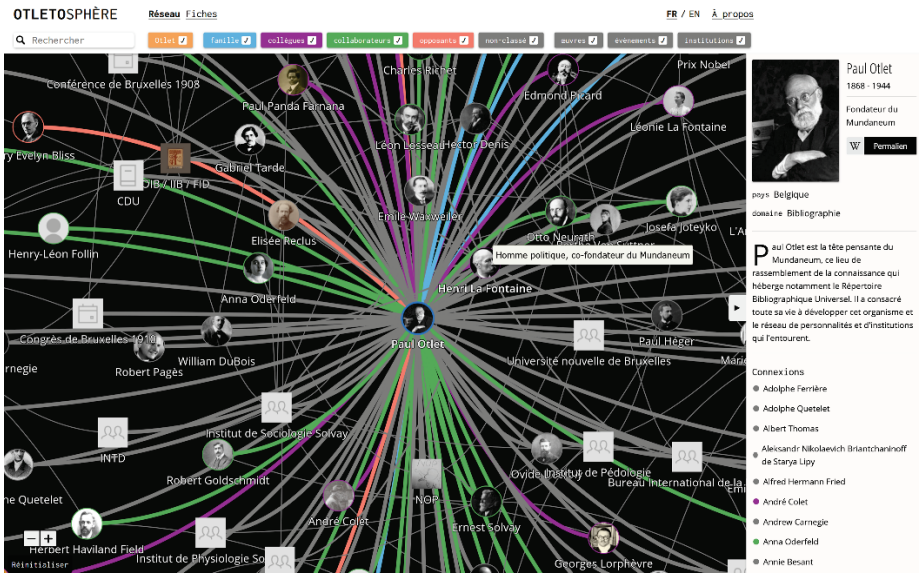

Figure 12. Exploration interface of the Otletosphère. Source: author's screenshot

Another area in which both Omeka Classic and Omeka S have been adopted is education. Their use extends beyond training information professionals in the conceptualisation, design, and implementation of repositories and digital libraries; they have also been employed in the development of a wide range of student and educational projects (Marsh, 2013; Cuenca & Kowaleski, 2018; Whalen, 2019; Gay, 2019; Laudito, 2020; Meunier, 2021), in collaborative initiatives between university libraries and students (Rinalducci, 2018; Beatty *et al.*, 2021), and for the management and dissemination of educational resources (Meunier, 2018; Meunier & Szoniecky, 2018).

The limitations previously noted concerning access to Omeka S resources via semantic web technologies have led some projects to develop additional integrations. One example is the Henri Poincaré archive (Bruneau et al., 2020; Lasolle, Bruneau & Lieber, 2020a; Lasolle *et al.*, 2021a), where improvements have been introduced to the information retrieval system for the corpus of letters studied. These include the use of lemmatisation to enhance search processing and retrieval results, as well as the incorporation of a SPARQL endpoint and a database based on Apache Jena to enable querying via these technologies (Figure 14). Although installed in a separate environment, this system is fully integrated into the website interface. The connection between the two systems is achieved through a dedicated script that converts the JSON-LD data available in Omeka S into Turtle format, which is then used to update Jena's RDF database. Similarly, Thalhath *et al.* (2021) have developed a dedicated module enabling access to Wikidata resources through its SPARQL API.

Figure 13. Excerpt from a virtual exhibition created by a student at Binghamton University (Gay, 2019). Source: author's screenshot

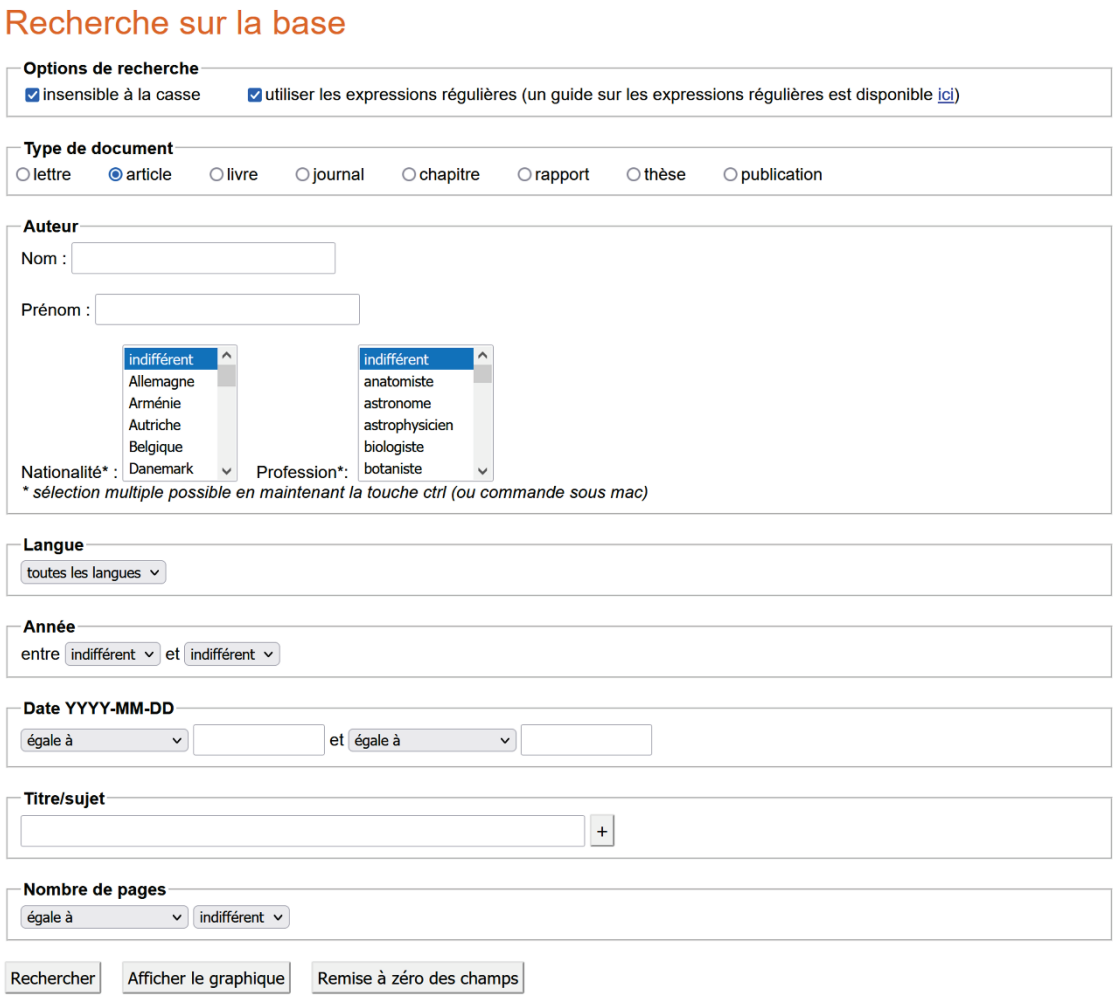

Figure 14. SPARQL query editor from the project La correspondance d'Henri Poincaré. Source: author's screenshot

3.7 Comparative Analysis

To select the applications used in the comparison with Omeka S, the following criteria were applied:

a) a content management system oriented towards the creation of digital repositories and libraries, offering advanced publishing functionalities;

b) free or open-source software;

c) a most recent update no more than one year old; and

d) managed or maintained by a reputable institution or demonstrating a substantial version history or level of maturity.

Based on these criteria, a total of five applications were selected for comparison: Omeka Classic, CollectiveAccess, Mukurtu, ResearchSpace, and Arches.

To document the characteristics of these applications, their official websites and documentation were consulted. When necessary, they were installed in a local or server environment or explored through publicly available demos. In recording functionalities, attention was paid both to the core features of each application and to the official modules or those endorsed by the project developers. Table 1 presents the results of the comparative analysis

Table 1. Comparison of Omeka S, Omeka Classic, CollectiveAccess, Mukurtu, ResearchSpace, and Arches. Source: author's own work

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
URL	https://omeka.org/s	https://omeka.org/classic/	https://www.collectiveaccess.org/	https://mukurtu.org/	https://researchspace.org/	https://www.archesproject.org/
Release date	2017	2008	2007	2012	2009	2013
Last stable version	3.0.2 (6/04/2021)	2.8 (25/032021)	1.7.11 (25/02/021)	2.1.8 (23/04/2021)	3.4.0 (25/05/2021)	5.1.1 (06/2021)
License	GPLv3	GPLv3	GPLv3	GPLv3	AGPLv3	AGPLv3
Developers	Corporation for Digital Scholarship	Corporation for Digital Scholarship	Whirl-i-Gig in collaboration with the Institute of Museum and Library Services.	Center for Digital Scholarship and Curation	The Andrew Mellon Foundation & British Museum	Getty Conservation Institute & World Monuments Fund
Technologies (programming language and database management system)	PHP / MySQL o MariaDB	PHP / MySQL	PHP / MySQL	PHP / MySQL o MariaDB	Java / MySQL	Python / PostgreSQL with PostGIS (Spatial and Geographic Objects)
Lightweight and coordinated workflows	No	No	Yes. Certain actions are customisable.	No	No	Customisable workflows with task tracking and publishing workflows.
User permissions management	Six user roles. Site-level permissions.	Four non-editable user roles.	Granular creation and management of different user roles and user groups.	Granular creation and management of different user roles and user groups.	Granular creation and management of different user roles and user groups.	Granular creation and management of different user roles and user groups.

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
Security	LDAP, login attempt limiter and log management, file validation, HTML filtering, and reCAPTCHA system.	File validation, HTML filtering, and reCAPTCHA system.	LDAP and support for other authentication methods.	Standard Drupal 7 authentication options.	LDAP, login attempt limiter, and log management.	No information available.
Open, flexible, and extensible system	Yes, an architecture based on a core plus modules. Oriented towards the creation of all types of projects	Yes, an architecture based on a core plus plugins. Oriented towards the creation of all types of projects.	No modules available, although the internal structure and part of the system's functionalities can be defined through XML configuration files. Oriented towards the creation of all types of projects.	Yes, an architecture based on a core plus modules. Oriented towards the creation of all types of projects.	Its internal architecture separates stored data and system logic from the implementation of third-party components used for presentation.	Yes, an architecture based on a core plus modules. Strongly oriented towards the management of immovable cultural heritage projects, with a significant geographical component.
Data integration and migration	Import formats: CSV Export formats: CSV Modules available for importing content from third-party applications.	Import formats: CSV and OAI-PMH Export formats: XML, JSON, CSV, METS, Atom, and RSS2 Plugins available for importing content from Flickr and YouTube.	Import formats: Excel, Google Spreadsheets, CSV, MARC21, FileMaker Pro, DSO XML, and InMagick XML Export formats: XML, MARC21, CSV, and PDF (finding aids)	Import formats: CSV Export formats: CSV, XML, and JSON	Native integration with Fedora Commons, OpenStreetMap, and Wikidata	Import formats: CSV, JSON-LD, GeoJSON, and KML Export formats: CSV, .shp, and JSON-LD Native integration with ArcGIS Pro

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
API	Yes	Yes	Yes	Yes	Yes	Yes
OAI-PMH	Yes	Yes	Yes	Drupal has an OAI-PMH module, but it has not been implemented in Mukurtu.	No	No
Suport	A relatively large and growing user community. Official forum and extensive documentation. Official hosting, maintenance, and customisation services. Available as a solution offered by several service providers.	A large and growing user community. Official forum and extensive documentation. Official hosting, maintenance, and customisation services. Available as a solution offered by a substantial number of service providers.	A not particularly large user community, mainly within the Anglophone sphere. Official forum, live chat, and extensive documentation. Hosting, maintenance, customisation, and migration services from other systems provided by Whirl-i-Gig. Available as a solution offered by several service providers.	A not particularly large user community. Official forum and extensive documentation. Limited number of related service providers. However, there are many service providers for Drupal, with which it shares underlying technologies.	A not particularly large user community. Adequate documentation. Very limited availability of technology providers.	A small user community. Official forum and extensive documentation. Official professional services provided by Farallon Geographics. Very limited availability of technology providers.
Interface customization	Yes, although it does not support multilingual sites. It does support multiple values in different languages within descriptions.	Yes, although it does not support multilingual sites, nor multiple languages for descriptive values.	Yes, it also supports multiple values in different languages within descriptions and multilingual sites.	Yes, it also supports multilingual sites.	Yes, and it is possible to create multilingual sites.	Yes, it also supports multiple values in different languages within descriptions and multilingual sites.

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
Digital Preservation tools	No	No	MD5 checksums, automatic creation of access copies. Integration with BagIt File Packaging Format Integration with Archivematica and Preservica (planned) Format migration and scheduled migration.	No	No	Change auditing.
File formats	Support for the main file formats.	Support for the main file formats.	Support for the main file formats.	Support for the main file formats.	Support for the main file formats.	Support for the main file formats.
Cataloguing templates	Yes, highly flexible (standards and data types).	Possibility of integrating different metadata standards (limited) and unrestricted creation of non-standard custom fields.	Yes, highly flexible (standards and data types).	Predefined, non-editable cataloguing template.	Yes, great flexibility in defining entities and description forms based on the ontologies used.	Yes, highly flexible (standards and data types).
Metadata	Any ontology-based vocabulary can be integrated.	DC, Qualified DC, PBCore, VRACore. Other schemas available through plugins requiring custom development.	Any XML-based metadata schema can be integrated.	DC and other non-standard descriptive elements	Ontologies compatible with the CIDOC Conceptual Reference Model and Schema.org	DC and the ability to import graph-based vocabularies for object description, as well as ontologies compatible with the CRM.

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
Controlled vocabularies	Integration with multiple linked data services (DBpedia, Geonames, VIAF, Wikidata, etc.). Module for creating controlled term lists.	Module for creating controlled term lists.	Integration with vocabularies from the Library of Congress or Getty vocabularies, among others.	Most fields use Drupal taxonomies to manage element values.	Integration of thesauri compatible with SKOS.	Integration of thesauri compatible with SKOS.
Search engine	MySQL and Solr.	MySQL and Elasticsearch.	MySQL and Elasticsearch.	MySQL and Elasticsearch.	SPARQL queries.	MySQL and Elasticsearch.
Simple search	Yes	Yes	Yes	Yes	Yes	Yes
Advanced search	Yes	Yes	Yes. Fully customisable.	Yes	Yes	Yes. Support for advanced searches based on thesauri.
User interaction and 2.0 social features	Social buttons, modules for crowdsourcing and collaborative transcription. Possibility of implementing forms for users to contribute to the repository.	Comments, social buttons, modules for crowdsourcing, collaborative transcription, annotations, social tagging, and favourites. Possibility of implementing forms for users to contribute to the repository.	Comments, social tagging, ratings, and social buttons. Possibility of implementing forms for users to contribute to the repository.	Comments.	No specific tools.	No specific tools.

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
Public interface customisation	Fully customisable. Six templates available in the official repository.	Fully customisable. Twelve templates available in the official repository.	Fully customisable. A single template available.	Fully customisable. A single template available.	Limited.	Fully customisable. A single template available.
Accessibility[2]	"Omeka strives to adhere to W3C web design standards and to be compliant with Section 508 of the Rehabilitation Act."	"Omeka strives to adhere to W3C web design standards and to be compliant with Section 508 of the Rehabilitation Act."	No information available.	No information available. Drupal does include a statement of intent: 'We are committed to ensuring that all core Drupal features adhere to W3C guidelines: WCAG 2.0 and ATAG 2.0.'	No information available.	WCAG 2.1 Level AA and selected Level AAA criteria.
SEO (Search Engine Optimization)	Support for semantic URLs on pages, but not on bibliographic records. XML sitemaps.	XML sitemaps.	No specific tools.	No specific tools.	No specific tools.	No specific tools.

[2] No expert assessment was carried out for the analysis of this criterion; instead, the information available in the documentation of each application was simply consulted.

Feature	Omeka S	Omeka Classic	CollectiveAccess	Mukurtu	ResearchSpace	Arches
Advanced publication	Integration with advanced image viewers, digital cartography, and timelines. Possibility of creating any type of resource such as virtual exhibitions or structured publications.	Integration with advanced image viewers, digital cartography, and timelines. Possibility of creating any type of resource such as virtual exhibitions or structured publications.	Integration with advanced image viewers. Image galleries, digital cartography, timelines, and virtual exhibitions.	Digital cartography	Integration with advanced image viewers. Data visualisations through statistical charts and relationship graphs.	Integration with advanced image viewers and a 3D object viewer. Integration with geographic information systems (GIS) such as Esri.

Beyond the information presented in the table, Omeka S users highlight features such as:

- its ease of installation, maintenance, and use (Whalen, 2019; Laudito, 2020; Markusic, 2020; Popović, Škorić & Rujević, 2021; Thalhath *et al.*, 2021);
- its robustness, particularly in comparison with Omeka Classic (Albrezzi, 2019; Jeanson *et al.*, 2020);
- the provision of an integrated environment for both content editing and publication (Lasolle *et al.*, 2020b);
- the rapid development of projects (Lombardo *et al.*, 2020b);
- the ease of developing a public interface (Lombardo *et al.*, 2020b);
- the integration of rich media resources (Karatas & Lombardo, 2020);
- the use of semantic web standards (Biay, Courtin & Marchesin, 2018; Neatrour & Myntti, 2019; Albrezzi, 2019; Avgousti, Papaioannou & Ribeiro, 2019; Lombardo *et al.*, 2020a; Lombardo *et al.*, 2021; Thalhath *et al.*, 2021), which allows repository objects to be related to other projects such as Wikidata or Geonames (Schweyer *et al.*, 2021);
- its interoperability (Bono-Lauriol, Dessaint & Queyroux, 2017; Benoît *et al.*, 2018; Jones & Muftic, 2020);
- its flexibility and editorial freedom (Popović, Škorić & Rujević, 2021; Biay, Courtin & Marchesin, 2018; Jones & Muftic, 2020);
- the ease of integrating and using standardised metadata schemas (Whalen, 2019; Dobre, Babeş & Bujor, 2020);
- its extensive catalogue of modules (Popović, Škorić & Rujević, 2021);
- its multisite architecture (Tzoc & Ubbes, 2017; Nishanbaev, Champion & McMeekin, 2019; Boyles, 2020; Tebeau, 2021); and
- its native integration with the Digital Public Library of America (Albrezzi, 2019).

Both Omeka S and other applications—such as CollectiveAccess or ResearchSpace—require a greater initial effort in developing the site architecture when compared with platforms like Omeka Classic (Li, 2020), Mukurtu, or Arches. This is largely due to the inherent flexibility of systems that offer extensive customisation options, which administrators must plan and configure before beginning to work with the content.

Omeka Classic remains an alternative that shares a substantial number of features with Omeka S. In comparison, Omeka S is technologically more advanced, although it does not yet offer the wide variety of plugins available for Omeka Classic, a far more established application. According to Berthereau (2021), the main module developer for both Omeka S and Omeka Classic, only 23% of the plugins available for Omeka Classic currently have an equivalent alternative in Omeka S. This weakness should, in principle, diminish over time, given the demonstrable commitment of the developers in previous projects. Even so, together with Omeka Classic and CollectiveAccess, Omeka S is among the solutions that provide the greatest number of content-presentation options beyond the traditional bibliographic record with access to the digitised document.

CollectiveAccess is characterised by its high degree of flexibility in selecting both the metadata standards to be used and the internal architecture and organisation of the application (Alcaraz, 2014). It also supports customised workflows and offers more extensive configuration of user permissions and roles. However, adding new functionalities is more complex than in modular applications such as Omeka S, for which a substantial set of pre-built modules is available. Finally, unlike Omeka, CollectiveAccess includes native support for creating multilingual sites.

Mukurtu is oriented towards the development of thematic repositories for Indigenous communities, although it can be adapted to other contexts (Shepard, 2014). It is the most restrictive application in terms of customising metadata element sets: users are limited to Dublin Core and a non-editable cataloguing template. It also offers fewer advanced publishing tools or social features. Nonetheless, it benefits from the stability and reliability of being a fork of Drupal, a robust and well-established CMS.

ResearchSpace, like Omeka S, focuses on semantic web technologies to provide a wide range of services and functionalities for projects ranging from digital collection management and digital humanities initiatives to the management and dissemination of scholarly publications. These functionalities centre on the creation, annotation, and argumentation of documents in various file formats, as well as on multiple forms of information visualisation, including document viewers, graphs, timelines, and statistical charts (Oldmann & Tanase, 2018).

Finally, Arches is a digital collection management system designed for the inventory and management of immovable cultural heritage (Myers *et al.*, 2012). It is the most specialised and least flexible of the five applications. Another of its key characteristics is its compatibility with different types of geospatial data, based on the standards and specifications of the Open Geospatial Consortium. This allows it to integrate a wide range of geographic information systems, such as ArcGIS, Google Earth, and Quantum GIS, as well as online services and APIs such as Google Maps and OpenStreetMap, among others (Myers, Dalgity & Avramides, 2016).

4. Conclusions

Omeka S is a young and evolving project that centres on semantic web technologies as an alternative for describing the information resources of a digital library or any other type of digital humanities initiative. Its status as free and open-source software, combined with its ease of installation, deployment, and use, makes it an ideal option for small and medium-sized institutions—typically those with more limited financial and staffing resources—as well as for educational contexts.

Its ability to manage a wide variety of websites or distinct projects within a single installation may be particularly valuable for smaller institutions that cannot afford to implement and maintain multiple applications yet need to launch several digital projects. In this sense, Omeka S helps democratise access to semantic web technologies across cultural institutions of all kinds. Nonetheless, its usefulness is not limited to smaller organisations; as the use cases reviewed demonstrate, it has also proved effective in large-scale projects.

Its considerable flexibility in selecting descriptive vocabularies, together with the growing number of modules that extend the core functionality of the application, makes Omeka S a compelling option regardless of the type of documentary collection or corpus on which a project is based. Its reliance on internationally recognised standards, and its capacity for integration with other applications—either through dedicated modules or via the OAI-PMH interoperability protocol—position it as a valuable component within digital ecosystems that already incorporate repository-management applications such as DSpace or Fedora Commons, which are comparatively less versatile and flexible in the presentation of content, or, in the latter case, lack a default public interface.

References

Albrezzi, Francesca (2019). *Virtual actualities: technology, museums, and immersion*. Doctoral Thesis. University of California. https://escholarship.org/uc/item/5tc2q2dt

Alcaraz-Martínez, Rubén (2012). Omeka: exposiciones virtuales y distribución de colecciones digitales. *BiD: Textos Universitaris de Biblioteconomia i Documentació, 28*. https://dx.doi.org/10.1344/105.000001829

Alcaraz-Martínez, Rubén (2014). CollectiveAccess, un sistema de gestión y difusión de colecciones de museos, archivos y bibliotecas. *Textos Universitaris de Biblioteconomia i Documentació, 33*. http://dx.doi.org/10.1344/BiD2014.33.23

Avgousti, Avgoustinos, Papaioannou, Georgios, & Ribeiro Gouveia, Feliz (2019). Content dissemination from small-scale museum and archival collections: Community reusable semantic metadata content models for digital humanities. *Code{4}lib Journal, 43*. https://journal.code4lib.org/articles/14054

Benoît, Pandolfi, Carmen, Brando, Eric, Mermet, Nicolas, Verdier, & Hermenault, Léa (2018). Oronce Fine: Une plateforme pour l'annotation sémantique de ressources cartographiques sur le Web de données. In *Conférence SAGEO 2018*. https://hal.science/hal-04310027v1

Beatty, Joshua F., Hartnett, Timothy C., Kimok, Debra, & McMahon, John (2021). 'And still we eise': Open pedagogy and black history at a rural comprehensive state college. In Davies, Kimberly, Cifton, Alexis. In *Open pedagogy approaches: faculty, library, and student collaborations*. http://hdl.handle.net/20.500.12648/1449

Berthereau, Daniel (2021). From Omeka Classic to Omeka Semantic. https://daniel-km.github.io/UpgradeToOmekaS

Biay, Sébastien, Courtin, Antoine, & Marchesin, Isabelle (2018). L'OMCI–Ontology of medieval christianity in images–de l'INHA: Une encyclopédie par l'image. *Archaeologia e calcolatori, 10*, 29–47. https://shs.hal.science/halshs-02990753v1

Bono-Lauriol, Sophie, Dessaint, Charlotte, & Queyroux, Nathalie (2017). ATELIER DIGIT_HUM 2016: les bibliothèques se livrent!. *Bulletin des bibliothèques de France, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB)*. https://www.archives-ouvertes.fr/ADH/hal-01794591v1

Boyles, Christina (2020). The Puerto Rico disaster archive: Preserving the cultural legacy of Puerto Rico. *Interdisciplinary Digital Engagement in Arts & Humanities (IDEAH), 1*(1), 1–13. https://doi.org/10.21428/f1f23564.c8ccde9c

Bruneau, Olivier, Lasolle, Nicolas, Lieber, Jean, Nauer, Emmanuel, Pavlova, Siyana, & Rollet, Laurent (2020). Applying and developing semantic web technologies for exploiting a corpus in history of science: the case study of the Henri Poincaré correspondence. *Semantic Web: Interoperability, Usability, Applicability*, 1–20. http://semantic-web-journal.net/system/files/swj2441.PDF

Cascón-Katchadourian, Jesús, Ruiz-Rodríguez, Antonio-Ángel, & Alberich-Pascual, Jordi (2018). Revisión, análisis y evaluación de sistemas para la gestión de activos multimedia en organizaciones. *Revista española de documentación científica, 41*(1). https://doi.org/10.3989/redc.2018.1.1481

Cuenca, Esther Liberman, & Kowaleski, Maryanne (2018). Omeka and other digital platforms for undergraduate research projects on the Middle Ages. *Digital medievalist, 11*(1), 3. http://doi.org/10.16995/dm.69

Dobre, Mihnea, Babeș, Ovidiu, & Bujor, Ioana (2020). Cartesian cosmological illustrations: A digital approach. In *Early Modern Cosmology Between «Mosaic Physics» and Mechanical Philosophy (1650–1713) – PN-III-P1-1.1-TE-2016-0710*. University of Bucharest. https://doi.org/10.5281/zenodo.4299767

Deuff, Olivier Le, Perret, Arthur, Borel, Clément, & David, Jean (2020). Surfer dans l'Otletosphère des outils pour visualiser et interroger le réseau de Paul Otlet. In *H2PTM'19, de l'hypertexte aux humanités numériques*, (pp. 65–76). https://archivesic.ccsd.cnrs.fr/sic_02480515

Gayol, Víctor, & Melo-Flórez, Antonio (2018). Jurisdictional culture and memory digitization of the 'Government of Justice': data modeling and digital approach for the legal history of Ibero-America. *Culture & History Digital Journal, 7*(2), e017. https://doi.org/10.3989/chdj.2018.017

Gay, Amy E. (2019). Using a content management system for student digital humanities projects: A pilot run. *Library Scholarship, 46*. https://orb.binghamton.edu/librarian_fac/46

Geslot, Jean-Charles, & Rebolledo-Dhuin, Viera (2020). Du livre imprimé au Web sémantique: Le projet du Dictionnaire des éditeurs français du XIXe siècle. *Humanités numériques, 2*. https://doi.org/10.4000/revuehn.426

Hoyle, Wendy (2017). How the smallest state in Australia ended up leading in the Australian GLAM sector: The power of collaboration when you only have a small project budget. In *HETA 2017 – Connecting Minds, Creating the Future*. http://dx.doi.org/10.4226/77/5a8366d486230

Jeanson, Loïc, Laroche, Florent, Kerouanton, Jean-Louis, & Bernard, Alain (2020). Knowledge management for modeled heritage objects, requirement specifications towards a tool for heterogeneity embracing. *International journal on interactive design and manufacturing, 14*, 1337–1345. https://doi.org/10.1007/s12008-020-00712-6

Jones, Kerry, & Muftic, Sanjin (2020). Endangered African languages featured in a digital collection: The case of the ǂKhomani San Hugh Brody Collection. In *Proceedings of the First workshop on Resources for African Indigenous Languages (RAIL)* (pp. 1–8). https://www.aclweb.org/anthology/2020.rail-1.1

Karatas, Tugce; Lombardo, Vincenzo (2020). Multiple perspective account of digital curation for cultural heritage: tasks, disciplines and institutions. In *UMAP '20 Adjunct: Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 325–332). https://doi.org/10.1145/3386392.3399277

Kelly, Jason M., & Horan, John (2020). Archive as pedagogy: oral history and a Journal of the plague year. *Collections: a journal for museum and archives professionals*. https://doi.org/10.1177%2F1550190620981029

Kelly, Jason M. (2020). The COVID-19 oral history project: some preliminary notes from the field. *The Oral History Review, 47*(2), 240–252. https://doi.org/10.1080/00940798.2020.1798257

Kubik, Tomaszm, & Kwiecień, Agnieszka (2020). Resolving dilemmas arising during design and implementation of digital repository of heterogenic scientific resources. *Applied Sciences, 11*(1), 215. https://doi.org/10.3390/app11010215

Lasolle, Nicolas, Bruneau, Olivier, & Lieber, Jean (2020a). Recherche d'informations dans la correspondance d'Henri Poincaré: outils et méthodes. In *Humanistica 2020*. https://hal.univ-lorraine.fr/hal-02570879

Lasolle, Nicolas, Bruneau, Olivier, Lieber, Jean, Nauer, Emmanuel, & Pavlova, Siyana (2020b). Assisting the RDF annotation of a digital humanities corpus using case-based reasoning. In *The 19th International Semantic Web Conference – ISWC 2020*. https://doi.org/10.1007/978-3-030-62466-8_38

Lasolle, Nicolas, Bruneau, Olivier, Lieber, Jean, Nauer, Emmanuel, & Pavlova, Siyana (2021a). Assister l'édition manuelle de données RDF à l'aide du raisonnement à partir de cas. In *32ème Journées Francophones d'Ingénierie des Connaissances (IC 2021)*. https://hal.univ-lorraine.fr/hal-03229501

Lasolle, Nicolas, Bruneau, Olivier, Lieber, Jean; Rollet, Laurent (2021b). Temporal knowledge representation for historical vorpora: Application to the Henri Poincaré correspondence corpus. In Semantic web. http://www.semantic-web-journal.net/content/temporal-knowledge-representation-historical-corpora-application-henri-poincaré

Laudito, Ilenia Eleonor (2020). Workflows, digital data management and curation in the RETOPEA project. In *Atti del IX Convegno Annuale dell'Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD). La svolta inevitabile: sfide e prospettive per l'Informatica Umanistica* (pp. 130–135). http://doi.org/10.6092/unibo/amsacta/6316

Leon, Sharon (2016). What to expect in Omeka S. In *Omeka S. Github repository*. https://github.com/omeka/omeka-s/wiki/What-to-expect-in-Omeka-S

Li, Jasmine (2020). Omeka Classic vs. Omeka.net. *Emerging Library & Information Perspectives, 3*(1), 232–236. https://doi.org/10.5206/elip.v3i1.8625

Lombardo Vincenzo, Damiano Rossana, & Karatas Tugce; Mattutino Claudio (2020a). Linking ontological classes and archaeological forms. In Pan J.Z. et al. (eds.). *The Semantic Web – ISWC 2020. ISWC 2020. Lecture Notes in Computer Science*, 12507. https://doi.org/10.1007/978-3-030-62466-8_43

Lombardo, Vicenzo, Karatas, Tugce, Daminao, Rossana, Mattutino, Claudio, & Sasakura, Mariko (2020b). Bringing digital curation to archaeological projects: evidence from the BeArchaeo project. In *Proceedings of AVI2 CH 2020: Workshop on Advanced Visual Interfaces and Interactions in Cultural Heritage*. http://hdl.handle.net/2318/1760490

Lombardo, Vicenzo, Karatas, Tugce, Gulmini, Monica, Guidorzi, Laura, & Angelici, Debora (2021). Transdisciplinary approach to archaeological investigations in a semantic web perspective. In *Semantic web*. http://www.semantic-web-journal.net/content/transdisciplinary-approach-archaeological-investigations-semantic-web-perspective

Lovejoy, Henry B. (2020). Who did what when? Acknowledging collaborative contributions in digital history projects. In *Esclavages & Postesclavages = Slaveries & Post-Slaveries, 4*. https://doi.org/10.4000/slaveries.2717

Markusic, Kirsten (2020). Investing in preservation: my role as a corporate historian at a Family-Owned Bank. *MAC Newsletter, 47*(3), 29–31. https://www.iastatedigitalpress.com/macnewsletter/article/id/12202

Marsh, Allison C. (2013). Omeka in the classroom: The challenges of teaching material culture in a digital world. *Literary and Linguistic Computing, vol. 28*(2), 279–282. https://doi.org/10.1093/llc/fqs068

Marsh, Emily (2017). Chickens, aprons, markets, and cans: How the National Agricultural Library uses Omeka as its content management system for digital exhibits. *Digital library perspectives, 33*(4), 361–377. http://dx.doi.org/10.1108/DLP-03-2017-0009

Mensaque-Urbano, Julia (2017). Expobus, un nuevo espacio virtual de exposiciones de la Biblioteca de la Universidad de Sevilla. *RUIDERAe: revista de unidades de información, 12*. https://ruidera.uclm.es/xmlui/handle/10578/16855

Meunier, Jean-Marc (2018). Valorisation of a pedagogical digital archive: The example of and ebook enriched from an ontology. In *Digital Tools & Uses Congress (DTUC'18), October 3-5, 2018*. https://doi.org/10.1145/3240117.3240138

Meunier, Jean-Marc; Szoniecky, Samuel (2018). An ontology-based enriched ebook for teaching statistics. In *ICOTS 10, 10th International Conference on Teaching Statistics*. https://hal-univ-paris8.archives-ouvertes.fr/hal-01841292

Meunier, Jean-Marc (2021). Une approche instrumentale de l'éducation aux débats socio-scientifiques en ligne. *Médiations et médiatisations, 5*, 199–209. https://doi.org/10.52358/mm.vi5.164

Myers, David; Dalgity, Alison, Avramides, Ioannis, & Wuthrich, Dennis (2012). Arches: An open source GIS for th e inventory and management of immovable cultural heritage. In Ioannides M., Fritsch D., Leissner J., Davies R., Remondino F., Caffo R. (eds). *Progress in Cultural Heritage Preservation. EuroMed 2012. Lecture Notes in Computer Science: Vol. 7616*. https://doi.org/10.1007/978-3-642-34234-9_87

Myers, David, Dalgity, Alison, & Avramides, Ioannis (2016). The Arches heritage inventory and management system: A platform for the heritage field. *Journal of cultural heritage management and sustainable development, 6*(2), 213–224. https://doi.org/10.1108/JCHMSD-02-2016-0010

Neatrour, Anna, & Myntti, Jeremy. (2019). Western Name Authority File: A pilot regional name authority project. *Journal of library metadata, 19*(1–2), 19–38. https://doi.org/10.1080/19386389.2019.1589685

Nishanbaev, Ikrom, Champion, Erik, & McMeekin, David A. (2019). A survey of geospatial semantic web for cultural heritage. *Heritage, 2*(2), 1471–1498. https://doi.org/10.3390/heritage2020093

Oldmann, Dominic, & Tanase, Diana (2018). Reshaping the knowledge graph by connecting researchers, data and practices in ResearchSpace. In Vrandečić, D. et al. (eds.). *The Semantic Web – ISWC 2018. ISWC 2018. Lecture Notes in Computer Science: Vol. 11137*. https://doi.org/10.1007/978-3-030-00668-6_20

Omeka (2021a). Sites management. In *Omeka S user manual*. https://omeka.org/s/docs/user-manual/sites

Omeka (2021b). Vocabularies. In *Omeka S user manual*. https://omeka.org/s/docs/user-manual/content/vocabularies/

Peñalver, Eduardo, & Bocanegra, Jesús (2018). Omeka S para exposiciones digitales. In *XVII Workshop REBIUN. Conectar tecnologías, enlazar contenidos, innovar servicios*. http://hdl.handle.net/20.500.11967/215

Popović, Petar, Škorić, Mihailo, & Rujević, Biljana (2021). The use of the Omeka Semantic platform for the development of the University of Belgrade, Faculty of Mining and Geology Digital Repository. Infotheca: *Journal for Digital Humanities, vol. 20*, 1–2. https://doi.org/10.18485/infotheca.2020.20.1_2.9

Rinalducci, Jenna (2018). Client/provider relationship: Library as client and art students as provider. In *Proceedings of the Sixth Conference for Entrepreneurial Librarians: Vol. 3* (pp. 23–31). http://libjournal.uncg.edu/pcel/article/view/1822

Schweyer, Anne-Valérie, Mermet,Eric, Paget, Adrien, Khounlivong, Ellie, & Noûs, Camille (2021). Contribution of the digital humanities to historical research in central Vietnam: GIS & semantic web annotations. https://hal.archives-ouvertes.fr/hal-03098418

Shepard, Michael (2014). Review of Mukurtu content management system. *Language Documentation & Conservation, 8*, 15–325. http://hdl.handle.net/10125/24610

Tebeau, Mark (2021). A Journal of the plague year: rapid-response archiving meets the pandemic. *Collections: A Journal for Museum and Archives Professionals*. https://doi.org/10.1177/1550190620986550

Thalhath, Nishad, Nagamori, Mitsuharu, Sakaguchi, Tetsuo, & Sugimoto, Shigeo (2021). Wikidata centric vocabularies and URIs for linking data in semantic web driven digital curation. In Garoufallou E., Ovalle-Perandones MA. (eds). *Metadata and Semantic Research. MTSR 2020. Communications in Computer and Information Science: Vol. 1355* (pp. 336–344). https://doi.org/10.1007/978-3-030-71903-6_31

Tzoc, Elias, & Ubbes, Valerie A. (2017). The Digital Literacy Partnership website: Promoting interdisciplinary scholarship between faculty, students, and librarian». *New Review of Academic Librarianship, 23*(2–3), 195–208. http://dx.doi.org/10.1080/13614533.2017.1333013

Whalen, Zach (2019). Teaching with objects: Individuating media archaeology in digital studies. *The Journal of Interactive Technology & Pedagogy, 15*. https://jitp.commons.gc.cuny.edu/teaching-with-objects-individuating-media-archaeology-in-digital-studies

Wittmann, Rachel, Neatrour, Anna, Cummings, Rebekah, & Myntti, Jeremy (2019). From digital library to open datasets: Embracing a 'collections as data' framework. *Information Technology and Libraries, 38*,(4), 49–61. https://doi.org/10.6017/ital.v38i4.11101