A Proposal in Creating a Semantic Repository for Digital 3D Replicas: The Case of Modernist Sculptures in Public Spaces of Rio De Janeiro

Daniela Lucas da Silva Lemos*, Dalton Lopes Martins**, Asla Medeiros e Sá***, Luciana Conrado Martins**** and Danielle do Carmo*****

*Library Science Department of the Federal University of Espírito Santo (UFES)
Av. Fernando Ferrari, 514, Vitória, ES, Brazil, < daniela.l.silva@ufes.br >
**Universidade de Brasilia, Faculty of Information Science,
Campus Darcy Ribeiro, Brasília, DF, Brazil, <daltonmartins@unb.br>
***Fundacio Getulio Vargas, Praia de Botafogo, 190 s 523, Rio de Janeiro RJ, Brazil, <asla.sa@fgv.br>
****Percebe, Epeira, 140, Vila Madalena, São Paulo, SP, Brazil, <lucianamartins@percebeeduca.com.br>
*****Universidade de Brasilia, Faculty of Information Science,
Campus Darcy Ribeiro, Brasilia DF, Brazil, <docarmo.danielle@gmail.com>

Daniela Lucas da Silva Lemos is Professor and Researcher at the Department of Library Science and the Postgraduate Program in Information Science at the Federal University of Espírito Santo, Vitória, Brazil. She holds a Bachelor in Information Systems Administration, a specialization in Strategic Information Management, a Masters Degree in Information Science and a PhD in Information Science. She has experience in the area of Information Science, with an emphasis on Information Organization, Knowledge Representation and Information Retrieval, acting mainly on the following topics: Knowledge Organization, Ontologies, Metadata Standards, Semantic Web, Linked Data and Semantic Annotation of Multimedia Resources on the Web. She has experience with IT, with an emphasis on Databases, Software Engineering and Analysis and Design of Information Systems.

Dalton Lopes Martins is Professor on the Librarianship and Information Science course and on the Postgraduate Program in Information Science PGGCinf at the Faculty of Information Science (FCI) at the University of Brasília (UnB). He coordinates the Tainacan research project (free software for the social construction of digital repositories) in partnership with the Brazilian Institute of Museums (IBRAM), the National Arts Foundation (FUNARTE) and the Brazilian Institute of National Historical and Artistic Heritage (IPHAN).

Asla Medeiros e Sá is a mathematician, Ph.D. in Sciences with an emphasis on Computer Graphics from the Institute for Pure and Applied Mathematics (Visgraf at IMPA), 2006. She has been Lecturer and Researcher at the School of Applied Mathematics at Getulio Vargas Foundation (FGV/EMAp) since 2007. From December 2015 until August 2016 and from July 2012 to February 2013, she has been in a post-doc position at the University of Brighton, UK, at the Cultural Informatics Research Group. There she became involved with additive manufacturing within the context of cultural heritage applications and developed research in cellular structures for additive manufacturing. Since then, she has organized seminars and led research projects applying Computer Graphics to the Cultural Heritage sector. In addition, in 2017, she was involved with research on periodic tilings.

Luciana Conrado Martins is a Historian, Master in Museology and PhD in Education from the University of São Paulo, who works mainly in museology, education and digital culture. Currently, she is director of the company Percebe, professor on the Postgraduate Program in Arts, Heritage and Museology at the Federal University of Piauí, professor of the specialization course in Museology, Culture and Education at the Pontifical Catholic University of São Paulo, and coordinates the network articulation of the Tainacan project (UnB).

Danielle do Carmo is a doctoral student on the Postgraduate Program in Information Science at the University of Brasília. Graduated in History, she is a specialist in creative processes and products and has a master's degree in Social Memory and Cultural Heritage. She has been investigating, since 2016, informational aspects related to the selection, organization and dissemination of digital objects from cultural and heritage collections through services and information systems on the internet.

Lemos, Daniela Lucas da Silva, Dalton Lopes Martins, Asla Medeiros e Sá, Luciana Conrado Martins and Danielle do Carmo. 2022. "A Proposal in Creating a Semantic Repository for Digital 3D Replicas: The Case of Modernist Sculptures in Public Spaces of Rio De Janeiro." *Knowledge Organization* 49(3): 151-171. 51 references. DOI:10.5771/0943-7444-2022-3-151.

Abstract: The demand for integrating and sharing heterogeneous data online has attracted the interest of cultural institutions in making information access and retrieval more effective via Semantic Web technologies. The present study proposes a digital repository for 3D scans of modernist sculptures in public spaces in the city of Rio de Janeiro, Brazil, with a view to ensuring access, use, reuse and preservation of this information. This is a qualitative exploratory experimental study based on the scientific literature and specific empirical material. It presents the analysis results of vocabularies for physical artifact documents and their digital counterparts on the Semantic Web and a discussion on how these align with the nature of the metadata determined here, as well as a metadata modeling prototype implemented on the Tainacan platform and aimed at cataloging digital 3D replicas. We claim that the proposed model for documenting cultural heritage assets on Tainacan is easy to implement, in that it uses accessible technology with a wide internet user base, highly expressive in its descriptions of 3D and multimedia content and based on well-established metadata and ontology standards recommended by regulatory bodies and communities such as the World Wide Web Consortium and International Organization for Standardization.

Received: 23 March 2021; Revised 25 July 2022; Accepted 26 July 2022

Keywords: networked heritage documents, digital repositories, 3D digitization, semantic annotation

1.0 Introduction

In recent decades, the creation and availability of online collections has grown considerably, with projects involving the digitization of unique icons or collections of cultural objects (Hildebrand et al. 2010; Pattuelli 2011; Potenziani et al. 2015; Rademaker et al. 2015; Scopigno et al. 2017; Potenziani et al. 2018), aimed at expanding the democratization of cultural knowledge on the internet in order to preserve, recover and conserve cultural heritage.

Projects such as 3DPetrie[1], STARC[2] and 3D-ICONS[3], among others, investigated the application of three-dimensional (3D) digital technologies in documenting digital 3D replicas related to cultural heritage. Scopigno et al. (2017) summarize the state of the art of the challenges faced by current digitization initiatives, emphasizing that despite the availability of sharing platforms, there are still gaps in structured approaches to generate and aggregate the results of a large number of 3D digitization initiatives. This includes the resulting digital 3D replicas as well as all the content produced during the digitization process. It is also important that information on the metadata used in these projects be shared between practical communities in order to share experiences on description standards for this type of resource.

In addition, the demand for integrating and sharing heterogeneous data online has attracted the interest of cultural institutions, such as libraries, archives, museums and documentation centers, among others, in making information access and retrieval more effective via Semantic Web technologies. Linked Open Data (LOD) is an initiative affiliated with the Semantic Web that provides a mechanism for linking data and metadata from different sources in order to generate new knowledge through semantically associated information resources (Bizer et al. 2009; Machado et al. 2019). To that end, communities involved in this initiative have developed languages and vocabularies based on standards recommended by the World Wide Web Consortium (W3C) in order to represent and organize their data in a global graph of interlinked data (Knoblock et al. 2017; Fink 2017; Gueguen et al. 2017; Charles et al. 2017).

However, society's use of these technologies raises questions regarding the existence of multiple formats, topics, languages, cultures and target audiences (lay communities and experts), leading to challenges in producing, organizing and disseminating information based on the interests of each type of user. These can be considered the main challenges for cultural heritage institutions interested in publicizing their collections online (Doerr and Iorizzo 2008; Hyvönen 2012). According to Doerr and Iorizzo (2008), there are two types of obstacles to promoting global networks of knowledge, namely social and technical. The social obstacle is that integration requires a commitment on the part of users to curate data and adhere to common standards, while the technical obstacle is the adoption of a generic information model to formalize a common concept for the cultural community, which would give rise to a multitude of research questions that could exploit advanced reasoning methods.

In this respect, this article discusses partial research results from a case study that utilized the maturity of 3D digital technologies and online data processing to give cultural heritage communities open access to documents involving cultural objects. The initial proposal involved 3D digitization of modernist sculptures displayed in open public spaces in the city of Rio de Janeiro, Brazil, with a view to creating a semantic repository for digital 3D replicas based on LOD principles. The project was conceived by the Schools of Social Sciences (FGV/CPDOC[4]) and Applied Mathematics (FGV/EMAp[5]) of the Getúlio Vargas Foundation (FGV), in conjunction with external collaborators, and is an interdisciplinary initiative that combines technology and computer graphics with expertise in historical documentation. The purpose of this semantic repository is to disseminate a detailed set of data aimed at different types of users (researchers, students, tourists and interested parties in general) to promote greater knowledge and use of modernist sculptures in Rio de Janeiro.

The project seeks to make a contribution to the challenge presented by Scopigno et al. (2017) by implementing a structured digital repository with different possibilities for information representation and technical functions to support data interoperability. In a way, a contribution is also made here to the challenges pointed out by Doerr and Iorizzo (2008) by providing a tool for users to catalog cultural information about 3D objects in a standardized way.

In order to systematically disseminate information via the repository from both a technical (documents pertaining to the files generated by the 3D creation process) and heritage perspective (documents related to the context of the sculpture), digital replicas must be cataloged in their locations of provenance and their original media and content aspects. In this respect, documentation on these cultural heritage assets would be obtained based on accessing and retrieving information that is relevant and interesting to content providers and their end users. As such, the present study sought to analyse potential semantic vocabularies, including ontologies, underlying conceptual models and metadata patterns, for the semantic modeling of metadata aimed at documenting digital 3D replicas that are the object of study of the project entitled "3D Digital Heritage: Modern Rio" (Patrimônio Digital 3D: O Rio Moderno in Portuguese). It is important to note that specifying the effective use of these vocabularies is a proposal for future research.

The aim of this paper is to describe the proposal of a semantic repository for 3D scans of public modernist sculptures in Rio de Janeiro, with a view to ensuring efficient access, use, reuse and preservation of the information. Considering that one of the greatest challenges is providing easily accessible digital tools for the public, technicians specialized in 3D digitization who need to document these processes and experts in producing conceptual documents for the digitized objects, we opted to adopt a flexible platform for metadata configuration and meaning attribution known as Tainacan (Martins et al. 2018), developed based on the WordPress content management system and used by a number of public and private cultural institutions in Brazil.

Although studies on delivering 3D models on the web are advanced (Scopigno et al. 2017), their integration with current semantic technology remains a challenge and requires further investigation. The use of ontologies is still an object of research, including semantic annotations for areas of 3D modeling, such as describing what a certain geometry represents (e.g.: number of vertices and faces, parametric relations). In this respect, the present study contributes significantly by mapping and analysing potential semantic vocabularies for recommendation in documenting digital 3D replicas in this first stage of the project. The use of metadata, both to describe the sculptures themselves and physically and semantically describe the multimedia files, is the recommended format for improving findability and ensuring adequate insertion of information resources into semantic networks such as LOD. Both technical attributes and high-level semantic characteristics can be included in metadata. Discussion about potential standards in line with the description of digital files involving 3D models that enrich documentation involving modernist works of art in public spaces in Rio de Janeiro is unprecedented and of great historical relevance in Brazil.

The article is organized as follows: this section contextualizes the proposal for documenting digital 3D replicas, presents the research challenges and outlines the objective of the article. Section 2 describes key concepts in the study and a review of use cases in cultural institutions interested in making their collections available within the linked open data paradigm. Section 3 describes the methodological aspects, Section 4 presents the results and discussions, and Section 5 contains the final considerations and direction of future research.

2.0 Organization and semantic representation of information

In the field of Information Science, particularly the area of knowledge organization (Dahlberg1993; Hjørland 2003; Hjørland 2007; Abbas 2010; Almeida 2013; Hjørland 2015; Hjørland 2016; Zeng and Qin 2016; Zeng 2019; Lemos and Souza 2020), research focuses on improving information retrieval systems, including studies on metadata standards, conceptual models, controlled vocabularies and ontologies for the descriptive and thematic treatment of documents in different media aimed at semantic integration and global availability of networked information resources (Knoblock et al. 2017; Fink 2018; Gueguen et al. 2017; Charles et al. 2017; Lemos and Souza 2020; Martins et al. 2022). In this respect, knowledge organization is linked to the analysis of concepts and relationships in a domain, leading to the synthesis of a knowledge organization system (KOS) to organize information resources, including: i) association, generating relationships; ii) representation, generating access points and indices for cataloging and indexing processes; iii) classification, promoting assignment and organization for documents; and iv) categorization, producing category schemes. Our study focuses on this field.

With respect to knowledge organization in digital environments, the advent of the internet paved the way for information specialists to improve methods for describing, organizing and retrieving remotely accessed digitized objects (Machado et al. 2019; Martins et al. 2022). In this new context of digital object production, organization and retrieval, working goals are not limited to creating symbolic representations of the permanent documents of a collection. They consist of creating new forms of hypermedia and so-called metadata, much of which can be extracted directly from the objects themselves (Gilliland 2016; Zeng and Qin 2016).

Therefore, the use of metadata can be considered a traditional method of adding semantics to information resources in processes such as description, indexing, classification, and cataloging within information retrieval systems (Zeng and Qin 2016).

Resource[6] in digital environments is generally described using annotation models expressed as tags, attributes and relations (Lemos and Souza 2020). Despite their widespread use on the web, annotation elements in the form of free-form natural language text are invariably subject to the semantic heterogeneity problem due to the ambiguous nature of natural languages, leading to issues such as polysemy, synonymy and specificity gaps (Foskett 1985; NISO 2005). This results in problems in information retrieval, such as: i) searching for isolated decontextualized words, making databases less visible to users and consequently, search mechanisms; ii) lack of context in the media items described (for example, how photos and videos are related to the text); iii) conceptual ambiguity (precise identification of the concept referred to); and iv) little relevance to the retrieved resource. These issues are generally addressed using controlled vocabularies to unambiguously identify resources or documents involved in information retrieval systems. Thus, users employ controlled vocabulary elements (e.g. terms, concepts) to descriptively and thematically represent resources and avoid ambiguities, and in search engines, user annotation terms are interpreted and linked to elements of a controlled vocabulary.

Ontologies can also be used as annotation models (Lemos and Souza 2020) in terms of the semantic treatment of data and metadata involved in the representation process. This makes it possible to describe and interlink resources through qualifiers, including concepts, instances, properties and restrictions, whose propositions are ensured by defining axioms. This model is recommended for the semantic annotation of documents, an approach based on semantic theory that gives an account of "meaning" whereby the logical connection of terms establishes interoperability between systems. As such, a set of open standards on data marking, modeling primitives and representation languages are promoted and maintained by the W3C to represent characteristics of digital objects on the web, including RDF (Resource Description Framework), RDF Schema and OWL (Ontology Web Language) (Hendler et al. 2020), as well as SPARQL (Simple Protocol and RDF Query Language), a query language for data modeled in RDF.

In this respect, linked data (Bizer et al. 2009) is a way of exposing and sharing data resources on the web and interlinking them with semantically related resources to make data available for consumption by humans and machines. Schandl et al. (2012) reported that part of the goal of the linked data paradigm is to connect different sources (exchanging semantically related information), forming a global graph of interlinked data that can be traversed by producers and consumers to discover new information. This vision is derived from the Linked Open Data project (Bizer et al. 2009; Machado et al. 2019), which identifies datasets available under open licenses, converts them into RDF triples according to linked data principles (by making each term in the vocabulary referenceable by a URI, for example) and publishing them on the web in the form of an interlinked "cloud". The content of this cloud is diverse and includes data on geographic locations, people, books, scientific publications, artwork, films, music, television, radio shows, genes, clinical trials, online communities, statistical data, and census results, among others. DBpedia[7] is an example of a KOS belonging to the Web of Data that displays information available on Wikipedia in a structured format, in addition to establishing links to other data sources.

The LOD project, therefore, proposes to annotate the content of digital objects or documents using domain ontologies (e.g. CIDOC CRM) or even controlled vocabularies (e.g. *Art & Architecture Thesaurus*) with less formal rigor (called Simple Knowledge Organization System, SKOS). Such representation of artifacts can be used as knowledge organization systems in the process of describing digital objects in various media in a web environment, each with its level of structural complexity, that is, by the formalism of their suitability for use in semantic technologies, especially in terms of relationships (Hjørland 2007). Thus, Shadbolt et al. (2006) consider semantic annotation as an underlying approach to the concepts advocated by the Semantic Web, providing an explanation of "meaning" whereby the logical connection of terms establishes interoperability between data and its applications.

Finally, a significant amount of data has been generated, connected and distributed on the web, especially multimedia data, which require efficient metadata to manage, organize and retrieve associated content. Benjamins et al. (2011) emphasize that important problems in the media sector are related to managing the textual and multimedia content explosion and that providing automatic support to manage this content requires conceptual capabilities of the solutions. According to Schandl et al. (2012), key enablers in successful integration of multimedia data lie in adopting ontologies as a formal means of describing their contents and technical characteristics. Several ontology initiatives aimed at the multimedia annotation domain are currently evolving in terms of maximizing semantic interoperability between data from producers and consumers of web content (Lemos and Souza 2020). However, some schemas available in the LOD network are insufficient for satisfactory semantic attribution since they do not comprise an adequate conceptual model to represent part of their realities. They also exhibit the following shortcomings in the quality of the information published: i) lack of conceptual description in the datasets; ii) absence of links in data schemes (to and between media); iii) lack of semantic expressiveness in data representation; and iv) no proper means to address and describe multimedia fragments.

Alternatives have been used to mitigate problems linked to semantic data treatment, including the use and reuse of domain ontologies aligned with (or oriented by) a generic or upper ontology (Guarino 1998; Borgo and Masolo 2009; Arp et al. 2015; Guizzardi 2005). The structure of an upper ontology aims to describe general concepts such as space, time, events, material, objects, etc., which are independent of a particular domain, and formalize them by clarifying the intended meaning of the terms adopted, preventing ambiguity and improving the quality of information representation.

The next section presents some of these ontology-driven conceptual models along with use cases in cultural institutions, the focus of this study.

2.1 Ontology-driven conceptual models for open data in cultural heritage

Interest in cognitive and conceptual models has grown in the field of Information Science (Almeida 2013; Lemos and Souza 2020) because problems with document and information representation in the digital context have become a challenge to interoperability between information systems and effective communication between users and information systems. It is also important to emphasize the great contribution that the adoption of conceptual models can impact the quality of data that is eventually used for training in applications and services using machine learning and artificial intelligence features. In the search for common interests, other areas and disciplines have joined information science, including cognitive science, computer science and recently, digital humanities (Koltay 2016; Clement and Carter 2017; Poole 2017), with a view to enriching methodological and theoretical resources and compiling a multidisciplinary body of knowledge capable of providing considerable scientific advances in information treatment. In this respect, contributions involving the initiatives and experiences described below aim to illustrate research efforts by information science and other communities in the area of conceptual modeling (concomitant to knowledge organization) to improve information retrieval.

According to Le Boeuf et al. (2018), institutions such as libraries, archives and museums have created pragmatic rules to describe collections and make them accessible to researchers, in addition to formats designed to index and store these descriptions on machines. These formats were guided by data models that sought to explain concepts and their relationships, but did not go beyond the data in their descriptions or focus on the semantic relationships in their representations.

From the 1990s, international researchers in the library and museum community (IFLA 2009; Le Boeuf et al. 2018) began developing conceptual models targeting the semantic aspects of their constructs and designed to provide a high-level view of the domain covered by bibliographic databases and intended for museums. Certain purposes were established in order to: i) serve as a valuable tool in assessing the relevance of existing rules for description, formats and data models in order to improve them; ii) convey a common concept that could be used to develop mediation tools between heterogeneous databases and ensure interoperability; and iii) serve as ontologies to contribute to the development of the Semantic Web paradigm. The International Committee for Documentation/Conceptual Reference Model (CIDOC CRM) falls within this framework (Le Boeuf et al. 2018).

The CIDOC CRM is a conceptual reference model intended to facilitate the online integration, mediation and exchange of heterogeneous cultural heritage information. It was developed by interdisciplinary teams from the fields of computer science, archaeology, museum curation, art history, natural history, library science, physics and philosophy, with the support of the CIDOC of the International Council of Museums (ICOM) and was accepted as a standard (ISO 21127) in September 2006.

Researchers involved in its design argue that the CIDOC is superior to most metadata models (whose resource is the central object of interest) in that it is event centric, whereby events can be related to different entities, such as actors, objects (physical and abstract), places and times (Doerr and Iorizzo 2008). For example, a historical image could be modeled as a network of chronological lines containing persistent items (objects and people) combined in events within a time period. This provides semantic enrichment of the data in terms of information retrieval because related (semantically interlinked) events can be collected, creating a powerful network of biographic and contextual data about people, documents, objects and places useful to, for example, scientific research. In short, the CIDOC ontology model allows all the items referenced in a resource document to be classified into formal categories, resulting in legible descriptions of events and objects.

The American Art Collaborative (AAC), a consortium of 14 institutions (13 museums and an archive) based in the United States[8], brought together a network of professionals and institutions around the idea of integrating and publishing the collection data of the institutions involved in LOD. Thus, the AAC is committed to creating a diverse critical mass of LOD on American art on the web, placing the collections of participating museums in the cloud and marking these data with LOD precepts. In 2014, the Linked.art conceptual model was developed, a simpler and more easily implemented adaptation of the CIDOC CRM model. Linked.art[9] is linked to the conceptual model recommended by the ICOM via the CIDOC and serves as a framework to facilitate the inclusion of global cultural institutions into the Semantic Web and linked open data.

In technical terms, Linked.art is an RDF profile of the CIDOC CRM that uses JavaScript Notation for Linked Data (JSON-LD)[10] and Getty Vocabularies[11] to describe cultural heritage objects. The model uses a subset of CIDOC CRM classes in conjunction with other RDF vocabularies to provide interoperable standards and models that can be interpreted both in JSON and RDF (Newbury 2018). According to Newbury (2018, 4) the model

> [...] focuses on usability and consistency, rather than completeness: as a design principle, it tries to cover 90% of the use cases of 90% of the organizations, with only 10% of the complexity of the full CRM ontology with all of its approved extensions.

Alexiev (2018, 21) states that there is currently no dominant and commonly accepted ontology for describing artworks and museum objects but considers Linked.art and CIDOC CRM as strong candidates. Below is a brief description of recent experiences such as Linked.art and CIDOC CRM in order to reinforce their applicability in potential projects focused on achieving interoperability between different types of collections of cultural institutions.

Knoblock et al. (2017) describe the experience and lessons learned by the American Art Collaborative in implementing 5-star linked data for the 14 institutions in the consortium. The project involved data mapping using the CIDOC CRM ontology, linking data with the Getty Union List of Artists Names (ULAN) vocabulary, applying tools to aid in these processes. The resulting data model later became known as Linked.art.

Fink (2018) discuses one of the products of the AAC project, namely a guide aimed at sharing with the museum community how the AAC approached linked open data, what tools were used, the trials encountered and lessons learned. The author also presents recommendations of good practices for museums interested in replicating his work. The document briefly addresses the construction of Linked.art and indicates that it is being applied to other datasets, such as the Getty Museum and Pharos (International Consortium of Photo Archives).

The principles and standards adopted in creating the Linked.art data model are described by Newbury (2018), highlighting the importance of data modeling that benefits different audiences, and the main critique of the CIDOC CRM, namely that its expansive scope and logical formalism make its practical application within computer systems a complex task for most software developers. This justifies the need for a more developer-friendly expression of the CIDOC CRM, achieved in the form of the Linked.art model.

Neely et al. (2019) address data use in museums and describe how visualization, design and art can be produced with data, providing new forms of access and understanding about museum collections. As an example, they cite the case of the small to medium-sized Georgia O'Keeffe museum that, after considering the benefits, opted to use the Linked.art model to publish data on their web-based collection. The authors also discuss the potential impact of openly available connected data on the scope of research about collections, providing examples of creative reuses for the available data.

Finally, the Europeana project is a comprehensive semantic repository that offers users free access to millions of books, paintings, museum collections and digitized archives from European cultural and scientific institutions (Charles et al. 2017). It uses a conceptual data model called the Europeana Data Model (EDM), similar to the CIDOC CRM in that it is event centric. Thus, the project aims to promote semantic mapping between different vocabularies, such as Friend-Of-A-Friend (FOAF) to describe people and organizations, the bibliographic ontology (BIBO) for describing bibliographic objects, Lightweight Information Describing Objects (LIDO) for museums, Encoded Archival Description (EAD) for archives, the Machine Readable Cataloging Record (MARC) for libraries, and Dublin Core for web resources in general. The EDM can be considered an ontology-driven conceptual model from the moment that it promotes cultural data enrichment through the high-level structure of the CIDOC. The Europeana project is deemed a global reference in aggregating cultural archives in LOD, ensuring their semantic interoperability.

As we have seen in this section, the trajectory of the data and conceptual models surveyed and discussed provides us with the realization that interoperability between information systems in the culture domain has always been sought by the communities involved. However, with the evolution of semantic web technologies for representing information resources, the semantic treatment of data has become much more sophisticated and attractive for applications, especially with the use of knowledge organization tools with formal rigor such as ontologies. In this field, foundational ontologies stand out in the sense that they support, by means of a formal and independent category system, domain ontologies, clarifying the intended meaning of the terms adopted by means of a set of semantic distinctions, avoiding ambiguity between concepts and relations, and improving, mainly, the quality of information representation in digital documents, making them more intelligent and therefore facilitating integration, alignment between schemas and vocabularies, and, finally, the data interpretation process by information retrieval systems. In this context, the modeling proposal of a semantic metadata scheme applied in this study stands out before the models raised and discussed in this section, including, for example, the ontologies in the cultural heritage domain CIDOC CRM and EDM, because the proposal of this research considers in its taxonomic structure, this research proposal considers in its taxonomic structure, foundational ontologies such as DOLCE and DUL (Borgo and Masolo 2009), which allow to formally describe from meta-categories based on cognitive, philosophical and linguistic aspects (event, object, agent, time, space, just to name a few) digital documents in various types of media in the context of cultural heritage.

The multimedia context, meanwhile, is modeled in ontologies selected for use and reuse from metadata typologies that characterize media resources, namely: content independent (metadata for management and administration of information resources), content dependent (visual metadata and for audio, both considered primitive level and, generally, their contents are extracted automatically by computational algorithms); content descriptive (metadata that associates media entities with real world entities); and provenance (metadata that describes the origin and production methods of the digital surrogate) for describing cultural objects in varied contexts and needs. The formal semantics, therefore, coming from the Semantic Web standards by which its structures are represented, in addition to providing means of transmission in some agreed syntax, ensures that the intended meaning of the semantics attached to the databases of cultural institutions can be shared between different applications, especially in projects aimed at integrating cultural collections in a linked open data environment.

3.0 Methodology

This study can be classified according to its research problem, objectives and technical procedures for data collection and analysis. Based on its approach to the problem, it can be considered qualitative, given the need to understand the little-known phenomenon investigated. As such, the determination, analysis and description of the methods and techniques suited to the research processes involved were possible by understanding this phenomenon, namely 3D digitization and its organized documentation published on a specific digital repository inherent to outdoor cultural heritage assets.

Based on its objective, this study can be classified as: i) exploratory, since it investigated and enhanced ideas about an emerging issue involving describing cultural and multimedia documents, including 3D models, which required a detailed analysis of the literature and specific cases in this regard; ii) descriptive, because it identified and described characteristics of the phenomenon studied based on the literature and specific empirical material; and iii) explanatory, given that based on the nature of the metadata identified in each stage of the proposed workflow for documenting digital 3D replicas, potential standards and vocabularies for describing these resources were indicated and the reasons for recommendation explained. With respect to the final point, an experimental method was used as support, since a repository was created and configured to establish some of these standards and vocabularies in order to propose a semantic metadata model for cataloging digital 3D replicas.

In relation to data collection and analysis procedures, the present study can be classified as bibliographic and documental because it used material published in the scientific literature as a source, including articles, conference annals, technical research reports, theses, dissertations, standards and documentary sources underlying the listed vocabularies.

The following subsections present the methodological stages of the study: 3.1 provides a general overview of the methodology proposed to identify, digitize and access the cultural heritage assets studied here; 3.2 presents the methods and techniques applied to identify, select and analyze the vocabularies recommended in each workflow proposed in 3.1; and 3.3 describes the technical procedures used in the mapping and prospection of technological solutions used as reference to compile a semantic repository as support for the digitization and creation of digital collections.

3.1 Documenting outdoor cultural heritage assets

The 3D Digital Heritage: Modern Rio project (Patrimônio Digital 3D: O Rio Moderno in Portuguese) is focused on documenting open and medium-large scale cultural heritage assets in the city of Rio de Janeiro. By open we mean both in the sense of their location in open spaces and the fact that they are openly accessible to the public, fostering a common sense of ownership and responsibility to care for them that relates to heritage conservation. The selected monuments and sculptures are from the Brazilian modernist period, an artistic movement expressed in architecture, landscape architecture, art, poetry, and literature. For almost a century, Brazilian modern and postmodern architecture and landscape architecture were world renowned. The focus on modern and contemporary history is therefore justified by the target institution, and the long-term objective is to interpret the choice and placement of public monuments in their historical context. Technology is the means by which the objective is achieved and not the objective itself.

The proposed methodology for documenting the outdoor cultural heritage assets is detailed in (Medeiros e Sá et al. 2019). The workflow consisted of three stages: i) identifying the heritage assets based on a combination of expert and community-led efforts; ii) producing the digital 3D replica; and iii) accessing the digitization results. Given that the purpose of this workflow was to document the entire process of digitizing outdoor cultural heritage assets, these stages are briefly described below to support the identification and description of the underlying metadata for each stage.

The process of identifying objects and their context metadata is partially collaborative. Participatory approaches can be used (Echavarria et al. 2018), whereby experts and community members join forces to establish what heritage, historical and present-day heritage are and how they are relevant to communities. To that end, as starting point we used the Inventory of Monuments in Rio Janeiro[12], a voluntary initiative that collected and published an inventory of public sculptures as well as interesting facts about several monuments based on historical sources. Our curators then selected a subset of sculptures of interest to define as priority for digitization based on several criteria, including technical restrictions, artistic relevance, context in the city and aesthetic considerations. The selected subset for the first phase of the project is illustrated in Figure 1, from left to right:

- Monumento à Juventude Brasileira (Monument to Brazilian Youth): Bruno Giorgi (1905-1993), a renowned Brazilian sculptor with Italian ancestry. Located at Gustavo Capanema Palace in the city center.
- Mulher (Woman): Adriana Janacópulos (1892-1978), a Brazilian sculptor of Greek descent. Located at Gustavo Capanema Palace in the city center.
- Evangelista Matheus (Matthew the Apostle): Alfredo Ceschiatti (1918- 1989), a Brazilian sculptor of Italian descent. Located in Catacumba Park in the city's South Zone.
- Estrutura (Structure): Sergio Camargo (1930-1990), the youngest artist in this initial set. Located in Catacumba Park in the city's South Zone.
- Background,Painel UFRJ (UFRJ Panel): Roberto Burle Marx (1909- 1994), renowned Brazilian landscape architect. Located in the School of Architecture building on the main campus of the Federal University of Rio de Janeiro, on the artificial island Ilha Fundão.

The second stage of the workflow, producing the digital 3D replica, was based on free and open-source software, including data acquisition, image processing, production and assessment of the digital 3D replica. For data acquisition, ter-

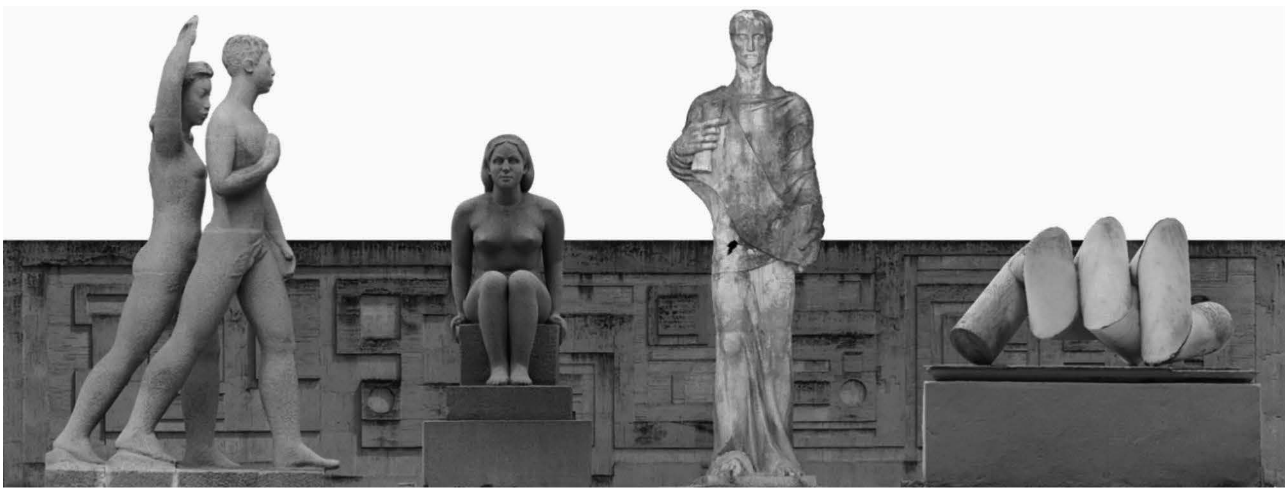

Figure 1. Set of monuments prioritized for digitization.

restrial and aerial photogrammetry were used (Gonizzi Barsanti and Guidi 2013), the latter only when necessary, using digital cameras to measure shape and texture and photometric properties to represent the shape and appearance of the digitized artifact. The data obtained were processed with photogrammetry software capable of automatically processing photograph datasets to produce a more accurate digital 3D replica. Since the 3D model was produced, it is important to assess its quality and document the entire digitization process. This ensures that both the process and its result are transparent to expert users who work with data handling.

The third and final stage, accessing the digitization results, allows the execution of post-processing phases that can be used to transform data into specific formats in order to provide access to digital 3D replicas through a variety of solutions (Potenziani et al. 2018), including web-based visualization, virtual reality platforms, orthophotographs and even the production of physical replicas via 3D printing. Additionally, from a technical standpoint, solutions can be adopted that combine digital information and physical elements of the environment, prompting discussion on the augmented reality context.

3.2 Identifying, selecting and analyzing vocabularies for the documentation

The methods and techniques used to identify, select and compile analysis categories involving a set of vocabularies were obtained from NeOn Methodology (Suárez-Figueroa et al. 2012), a scenario-based guide with a series of flexible steps to develop and reuse semantic vocabularies for the web. The Neon Methodology was chosen because it is a current methodological guide tested and validated in different domains, especially multimedia annotation, provides LOD initiatives and comes from frameworks that are widely accepted in advanced fields such as software and knowledge engineering. These guidelines were used to identify and select semantic vocabularies aimed at annotating cultural heritage documents with multimedia features (particularly 3D) by searching the literature and semantic web repositories.

The five methodological procedures followed are described below, based on some of the scenarios proposed in the guide: i) identify knowledge resources in reliable sources, including multimedia ontologies, ontology-driven conceptual models and cultural heritage metadata standards; ii) assess these resources according to recommended criteria such as coverage, accuracy and consensus on the knowledge and terminology used in the resource; iii) select the most appropriate vocabularies for the candidate resources assessed; iv) conduct a detailed analysis of the knowledge resources obtained and identify their underlying elements to create representations on different levels of abstraction (functional requirements and conceptual schemas, among others); and v) represent the resources in schemas to organize the knowledge acquired during analysis.

Initially, a domain study was conducted via documentary sources, including standards, articles and libraries of XML schemas related to metadata standards committed to the linked open data paradigm and aimed at describing artworks, that is, documents from the cultural heritage domain. The Dublin Core, VRA Core and LIDO metadata standards were chosen as reference material and to acquire knowledge of the domain.

Dublin Core[13] was developed in the 1990s when the internet was growing. A discussion at the 2nd World Wide Web Conference in Chicago in October 1994 identified the need for an infrastructure to enable online resource discovery. The Dublin Core Metadata Element Set (DCMES) became an international standard in 2003 (ISO 15386). Efforts to integrate Dublin Core with RDF language are currently becoming a reality. The Dublin Core Metadata Initiative (DCMI) and Semantic Web communities focus on converting metadata format into metadata vocabulary, that is, a collection of carefully defined properties to create descriptive sentences about resources (also known as namespaces). These changes introduced the notion of the Qualified Dublin Core to refine the resources described.

The VRA Core[14] was first developed in 1996. The current version 4.210, released in 2007, is expressed as an XML schema to support VRA Core interoperability and registration exchange. It is an internationally recognized metadata standard used as both an independent format and approved extension schema for the METS[15] (Metadata Encoding and Transmission Standard) for cultural heritage objects such as paintings, drawings, sculptures, architecture and photographs, among others. The VRA Core Oversight Committee [16] (Core OC) is a current initiative to produce an RDF ontology based on the XML Schema version that enables descriptions to be shared and interact with other web-based linked data resources, allowing users to easily find relevant images of cultural heritage objects, locations and subjects.

LIDO[17] is an XML harvesting schema developed in 2008 from several existing metadata standards. It was the result of a joint effort by international communities, including the CIDOC, to create a common standard that provides information and content on cultural heritage for multilingual portals and other aggregation repositories. It also sought to address what communities identified as shortcomings in Dublin Core when used to describe cultural materials and their digital substitutes.

Ontological resources for the multimedia annotation domain were selected based on an in-depth study presented by Lemos and Souza (2020), who compared and classified nine proposals based on the ISO MPEG-7[18] and Dublin Core. The study identified relevant characteristics that should be included in these ontologies to ensure multimedia resource interoperability and retrieval in a semantic network. The result of the ranking of candidate ontologies for reuse obtained in the comparative analysis identified the most prominent ontologies for multimedia annotation, in the following order: Media Ontology, M3O, COMM and M3 Multimedia.

Media ontology was proposed in 2009 by members of the W3C Media Annotation Working Group, which aims to improve interoperability between metadata schemas for web-based media resources, such as video, audio and images. The ontology was compiled using ontology engineering standards aimed at defining a set of key annotation properties to describe multimedia content, along with a set of mappings between the main metadata formats currently in use.

The Multimedia Metadata Ontology (M3O) was created in 2010 as a comprehensive model to represent metadata that describe multimedia, including combinations of metadata standards and models to semantically describe multimedia document presentations.

The Core Ontology for Multimedia (COMM) was developed in 2007 by a group of renowned researchers in the fields of multimedia, digital libraries and the Semantic Web. The main aim of COMM is to provide a fundamental conceptualization for multimedia description, generally covering a specific domain that deals with this type of content. Recognizing the semantic limitations of MPEG-7 but taking into account that the standard is a well-established multimedia knowledge base in the community, the COMM developmental team re-engineered it to create a formal representation of MPEG-7 descriptors with the same terminology convention.

M3 Multimedia was created in 2012 as part of a comprehensive ontology (addressing different domains and languages) denominated the M3 Ontology Network[19], the product of a Spanish research project involving entities such as the Ontology Engineering Group (OEG).[20] The Buscamedia project aimed to create a semantic search engine for multimedia resources with a view to achieving progress in the areas of semantics, audiovisual production and media distribution. The team that developed M3 included well-known researchers from the Facultad de Informática da Universidad Politecnica de Madrid and W3C Media Annotations Working Group.

Finally, ontology-driven conceptual models for the cultural heritage with coverage for provenance scenarios were selected based on a review of use cases in the literature involving the description of 3D cultural objects, namely the Linked.art in Section 2.1 and CRMdig models. CRMdig is an ontology-driven conceptual model designed to semantically organize provenance metadata (Doerr et al. 2016). It is an extension of the CIDOC CRM ontology, capable of capturing modeling and query requirements related to the provenance of digital objects (including production methods and stages) for a domain, especially cultural heritage.

The results of the vocabularies analysed were organized by establishing analysis categories (shown in Table 1) based on use cases for different projects involving criteria for the development and reuse of semantic vocabularies described in the NeOn guide, as well as specific categories for metadata research.

The first category, "nature of the multimedia metadata and the artwork itself", requires elucidation because it involves characteristics of metadata used to describe an actual physical work of art and its digital forms. This includes metadata about the types of media files and their 3D objects, metadata that describe the digital 3D object and the actual artifacts. As such, the elements of the vocabularies analyzed were organized into three metadata categories (supported by the literature), as follows: content independent, content dependent and descriptive content metadata.

Nature of the multimedia metadata and the artwork itself	Characteristics of the artwork and the multimedia context covered by the vocabularies analyzed, thereby supporting the use and reuse process.
Representation language	Verifying the language that represents the components (classes, properties, instances, axioms, etc.) of the vocabularies analyzed.
Code clarity	How easy the code is to understand; if the entities in the structure follow a pattern, are clear and coherent; whether there are comments; and if the code is well documented.
Knowledge extraction suitability	Related to how easy it is to identify and extract parts of knowledge from the vocabularies analysed.
Naming convention suitability	Verifying the rules associated with naming (terminology) the components.
Annotations in the terminology	The existence and quality of the annotations in the terminology elements of the vocabularies analysed.
Axioms in the terminology	The existence of axioms in the elements of the ontologies analysed, thus ensuring restricted interpretations.

Table 1. Analysis categories applied to the vocabularies.

The "content independent metadata" category addresses the management and administration of information resources and was grouped into four subcategories: i) creating and producing the resource; ii) classifying the resource; iii) information about the resource; and iv) using the resource.

Subcategory (i) contains characteristics that involve creating the artwork or media content and the resources associated with it; (ii) consists of characteristics aimed at classifying the artwork or multimedia resources, such as genre, subject, purpose, language, age classification, parental guidelines and subjective assessment; in (iii) the characteristics are related to the artwork or media storage, including format, compression and coding of the multimedia content; and in (iv) they reflect the rights of use, registration, availability and financial information about the artwork or multimedia content.

"Content dependent metadata" was grouped under the visual metadata subcategory, whose content is generally extracted automatically by computer algorithms. Visual metadata cover characteristics such as color, texture, shape, movement, location of temporal space regions and facial (or object) recognition in media content.

"Descriptive content metadata" associates entities within the content of the artwork or media content with those in the real world and were grouped into the following subcategories: i) media segments; ii) content semantics; and iii) personalized content. Subcategory (i) is related to content structure in terms of multimedia segments; (ii) involves objects, event and notions of the real world that can be abstracted from the artwork or multimedia content; and (iii) corresponds to personalized multimedia content formats to facilitate navigation, access and user interaction when consuming the content.

In considering uses with greater social emphasis of the proposed representation of information around these three metadata categories that have been proposed here, it is worth considering how these metadata categories can be used to represent social, economic, and possibly political information about the artworks and their digital counterpart. By allowing an expanded use and reuse of this information for different types and potential curatorial practices, it is understood that the proposal extends its social value, which is understood to be one of the central goals of the documentation of cultural objects. In this way, it is understood that the metadata from the "descriptive content" category can be used to describe social, historical, political and economic facts related to a given cultural object. In addition, it is understood that this metadata could describe potential educational uses of the work by proposing, for example, an educational activity that could be done and present a complete script for mediation with students. The fact that this category involves events, facts, objects, and general real-world information that relate to a cultural object allows it to expand its capacity for abstraction and build potential uses of information that meet initiatives of extroversion of social values that are desired to be promoted by the curators of the collection.

3.3 Mapping and prospection of technologies to organize digital collections

Producing a digital environment to organize and manage information resources that support researchers in producing, processing and interpreting digital 3D objects and disseminate information to a wide audience requires continuous analysis of the best technological solutions, international standards and connections with social media to determine the information needs of users. In this respect, it is vital for the project to continue mapping potential solutions and prioritize the free format in order to identify the potential, weaknesses and social aspects that should be considered when proposing the repository architecture.

The present study adopted the Tainacan digital repository[21] as a technological solution to organize and manage digital 3D collections. Given the previously mentioned characteristics of this repository, it is considered the best solution for the project objectives to insert, store, handle and visualize 3D models, their objects and related metadata, allowing efficient access, use, reuse and preservation of the information and ensuring the semantic integrity of the data in the LOD network.

The methodology enables continuous modeling of the information environment and a customized prototype that meets all the previously described requirements. The modeling includes a graphical interface, content classification formats, information retrieval modes, social interaction possibilities and an operational network for federated repositories. The stages defined were:

- identifying and assessing the information organization and documentation practices and conditions to define metadata (Section 3.1);
- analytical study to propose a semantic conceptual model for description and representation of the collection (Section 3.2);
- technical treatment of existing documentation, involving collection, analysis, normalization, cataloging, enrichment and availability on the Tainacan digital repository platform (this section).

The most important aspect of deciding to adopt Tainacan was the fact that the platform does not follow a rigid metadata standard for cataloging, allowing researchers to model their own set of metadata and, primarily, define a semantic URI to model each metadata element, attributing meaning and preventing ambiguity when interpreting the concepts used. This results in an easily accessible tool that enables product information to be exported from the catalog to generate RDF models in the formats adopted for the study.

4.0 Results and discussion

Section 4.1 presents the analysis results of vocabularies for physical artifact documents and their digital counterparts on the Semantic Web and a discussion on how these align with the nature of the metadata determined here. Section 4.2 presents a metadata modeling prototype implemented on Tainacan based on some of the vocabularies analysed to catalog the digital 3D replicas produced for the project.

4.1 Vocabulary analysis: processing digital 3D replica documentation

The vocabularies were selected for use and reuse in projects involving information organization in the cultural heritage domain, particularly digital 3D replicas. This justifies the proposed use in each stage of the workflow outline in the project scope.

Stage 1, identifying the heritage assets requires describing the artifact itself, depicted by the 3D model and referred to here as the Real World Object (RWO). In general, this object will be a public sculpture or monument in the city of Rio de Janeiro, but could also be a painting in a museum catalog. Thus, classification information would be useful for the semantic aspects involved (descriptive content metadata). Content independent metadata are also needed to describe the author of the artwork, as well as cultural context, geographic location, creation data, information sources, the title, historical period, material and techniques used in production, and author's rights, among others.

Stage 2, producing the digital 3D replica, requires metadata on the media object, referred to here as the Digital 3D Replica (3DR), which contain information on the file format, size, location and necessary software. In this case the metadata depend on the technique used to generate the 3D model, which in our research was photogrammetry. Media and 3D objects are digital files that represent a sculpture, building, design, etc., and therefore require information such as format, number of geometric objects and scale for model composition and visualization. Additionally, a set of visual content dependent metadata (color, texture, shape, among others) can be obtained automatically by processing the photograph data involved in producing the digital replica.

Stage 3, accessing the digitization results, demands metadata to describe the 3D Object (3DO) after processing for different purposes, including information about managing and personalizing this type of resource and portraying semantic aspects, such as annotations on regions of interest, objects, places, events and periods within the context of the replica. These aspects should be linked to the specific semantic vocabularies (or ontologies) that depict the entities involved.

The Dublin Core, VRA Core and LIDO metadata standards provide satisfactory coverage for the content independent category and can be used in all three workflow stages to describe the RWO, 3DR and 3DO. However, analysis of their structures showed no coverage aspects for content dependent metadata. It is important to note that both Dublin Core and VRA have qualifiers that expand the coverage of their elements, which positively affects the model extensibility. Additionally, all three standards have extensions for W3C representation languages, such as RDF, to ensure better insertion within the context and compliance with the interoperability demands of the Semantic Web and LOD. In the case of Dublin Core and LIDO, because a consolidated version already exists in RDFS, which is not yet the case for VRA, real entities represented in the media content can be linked to the semantic vocabulary (ontologies and the SKOS) of specific domains on the web. As such, with regard to the descriptive metadata needed in stages 1 and 3 of the workflow process, the semantic content can be covered by Dublin Core and LIDO.

The LIDO metadata standard seems to be the best suited to cataloging artworks because it was designed to deal with the digital content of cultural institutions, particularly museums, and is related to CIDOC CRM (Pitzalis et al. 2011). As a result, in conjunction with the cultural heritage ontologies discussed here (Linked.art and CRMdig), this standard can support the process of describing modernist sculptures (workflow first stage) and their digital counterparts (remaining stages). The set of elements compromising its schema provides a category structure for organizing administrative and descriptive metadata, including classifying objects, author's rights, relationships with other objects, and events the object participated in, among others. However, annotation supported by this standard will need to be complemented with other vocabularies (cited in this study) in order to cover specific aspects involving objects portrayed using three-dimensional techniques and documentation of their forms of production, especially for the types of metadata required for documentation in workflow stage 2.

The Media Ontology is recommended for content independent metadata, which are particularly useful in workflow stages involving 3DR and 3DO documentation. This is justified by the fact this ontology has a satisfactory coverage rate in relation to COMM, as reported by Lemos and Souza (2020), whereas M3 Multimedia reuses the Media Ontology for this metadata category.

COMM and M3 Multimedia offers descriptors suited to content dependent metadata, especially useful in workflow stage 2 for data processing to automatically generate metadata and, consequently, the 3D model. Both exhibit very similar visual coverage rates, particularly for descriptors involving color, texture, shape, and location of regions of interest. Metadata to describe 3D characteristics are present in both ontologies, primarily for shape-related visual aspects, since both are based on MPEG-7 for multimedia content description. This standard includes descriptors that may be useful in the present project because they cover 3D object characteristics such as symmetry, circularity, axis location, size and orientation of consecutive border segments, points of curvature and angles of curves.

M3O and COMM offer organized descriptors based on an upper ontology and multimedia design patterns, in addition to addressing semantic differences between media content and realization, making them useful for semantic organization of descriptive metadata and therefore applicable in stages 1 and 3 of the documentation process. Its architecture is modeled on an upper ontology (DOLCE for COMM; and DOLCE+DnS Ultralight – DUL for M3O) and the following three design patterns: Descriptions and Situations (D&S), Information and Realization Pattern and Data Value Pattern. The multimedia standards of both include the Annotation and Decomposition Patterns. In addition, multimedia standards operate under the semantics specified in the Information and Realization pattern, which represents the distinction between information objects and information realizations, whose separation is relevant in terms of providing a clear distinction between semantics (content of the media message) and data (media file format). Thus, annotations and decompositions may involve information objects and realizations. Metadata focused on content semantics are generally linked to instances of domain ontologies whose semantic labeling is organized within the taxonomy of an upper ontology. Both are part of a generic ontology that plays the role of organizing semantic labels from ontologies of specific domains into entities such as vent, object, time, place, etc., in addition to dealing with their relationships.

In regard to data provenance, the taxonomy of CRMdig contains representative classes of objects and events (as well as other interactive entities, such as person, place, time, etc.) involved in the production of digital artifacts, including devices, computer algorithms and people who participate in the workflow. Given the high level of the categories represented in the taxonomy of the model, metadata related to the 3D replica production cycle events can be structured based on five main issues:

i) the people or organizations who participated in the event;
ii) the location of the event;
iii) its date and time;
iv) the entities involved; and
v) the type of process and techniques applied.

This information is vital to the treatment of digital replicas of cultural objects in terms of guaranteeing a transparent relationship between the digital replica and real physical object, and allowing repeatability and verifiability.

Some event and object classes of CRMdig that could be applicable to the production stage of the digital 3D replica are:

i. D2 Digitization Process: transition from material object to digital representation;
ii. D7 Digital Machine Event: creation of the digital object with a device operated by a human actor;
iii. D8 Digital Device: information on the device used in replica production, such as scanners and cameras, etc.;
iv. D10 Software Execution: Information on the series of computer operations involved in producing the digital 3D replica;
v. D14 Software: Information of the software codes, computer programs, procedures and functions used to produce the digital 3D replica; and
vi. D21 Person Name: the name of the person involved in replica digitization.

Content dependent metadata related to the digitization of the sculptures could be aligned with the CRMdig D9 Data Object class, which represents the direct result of a digital measurement, containing quantitative properties related to physical entities. Descriptive metadata for specific annotations, that is, depending on the purpose of visualizing and accessing the replicas (workflow stage 3), could be organized into the following CRMdig event and object classes:

i. D3 Formal Derivation: the events that resulted in the creation of versions of the digital 3D replica based on what was produced, whereby these versions share representative properties with the original object, preserving the representation of some features but in a different way;
ii. D29 Annotation Object: the objects that describe the digital 3D replica, such as regions of interest, objects, places, events and periods involved in the context;
iii. D30 Annotation Event: the events that created the annotation object; and
iv. D35 Area (segmentation): information on the spatial location of a part of interest on the 3D model.

The fact that CRMdig is an extension of CIDOC CRM, which also proposes a simplified data model in the context of linked open data (Linked.art ontology), means the nature of the previously discussed metadata can be structured and linked into generic classes, such as person, object, activity, place and time, given the common conceptualization of the three ontology-driven conceptual models.

Finally, the metadata standards, multimedia ontologies and ontology-driven conceptual models analysed here cover the functional and nonfunctional requirements (Silva 2014) deemed important to future proposals for a comprehensive metadata modeling architecture to represent documents in the cultural heritage domain. Some of these requirements are listed below.

- Considers media content and realization in different formats, such as audio, image, text, 3D models and video: separating objects and their realizations is important because content independent metadata such as file size or media location on the web are typically applied to information realization, whereas descriptive metadata for multimedia content aim to describe the message to be conveyed to the consumer. As such, this separation is relevant in that it provides a clear distinction between content semantics and the media resource.
- Covers content independent, dependent and descriptive metadata: the vocabularies analyzed show reasonable coverage rates for multimedia content in the cultural heritage context.
- Covers metadata to describe the real physical resource, digital resource and its provenance data: schemas such as LIDO and the CRMdig and Linked.art models can establish integration through the common taxonomy provided by the terminology and its core ontology (CIDOC CRM). Thus, events associated with the production of the artwork and digital 3D replicas can be aggregated by semantic entities such as people, objects, places and times, contextualizing heritage documentation and favoring its retrieval.
- Uses an upper ontology as reference: semantically benefits the core taxonomy of the domain ontology by clarifying the intended meaning of the terms, supporting, for example, the integration of instances of media content with ontologies of specific domains.
- Based on extended multimedia standards of ontology design patterns: mitigates the challenges of reuse with acceptable and memorable diagrammatic visualizations for a specific set of competence issues (problem and its solution).
- Ensures semantic interoperability in relation to multimedia content on the web: guarantees that the intended meaning of the captured semantics can be shared between different LOD applications.
- Architecture allows separation of interests: the vocabularies analyzed enable clear separation of interests in relation to media, as follows: semantics of cultural heritage content, knowledge related to information resource management, structure and features of the multimedia content.

Table 2 summarizes the analysis results of vocabularies to guide use and reuse decisions regarding their structural aspects, based on predetermined functional and non-functional requirements.

4.2 Structure of the Tainacan repository for metadata modeling

Tainacan was used to model a cataloging form to create the digital repository for the project. The present study adopted the metadata model suggested by Pitzalis et al. (2011) and proposed using metadata created in the LIDO to be mapped and represented based on the CRMdig ontology model. The aforementioned study was highly flexible and easily adjustable for application in a digital repository that allows dynamic metadata configuration, making it easier for the metadata proposed by the LIDO model to be adopted without major technical difficulties in implementation. Since the metadata were created in the repository, the model proposes mapping entities, relationships and attributes for CRMdig, making it possible to export metadata from a system and computationally make it available in a semantic model in RDF. This is considered an important step in making it easier to adopt this type of standard for cultural heritage documentation, ensuring that the institution and professional have fewer concerns with standards that are easier to learn and implement, but more compatible with highly expressive ontologies (CIDOC CRM, for example) and allowing the metadata to be more widely reused.

Regardless of the name attributed to facilitate understanding by repository catalogers, a semantic URI was attributed to each metadata element according to the model used. For example, Figure 2 shows the metadata configuration referred to in the platform as Digitization Date (Data da Digitalização da Obra), representing the CRMdig entity E61 Time Primitive in the model. Attribution of the semantic URI allows the meaning of each metadata element to be identified without ambiguity.

Since the metadata from the LIDO standard were configured on the Tainacan platform based on the technical metadata creation functions of the software and attributed to CRMdig semantic identifiers, as suggested by Pitzalis et al. (2011), the comprehensive specifications and exhibition of a digital object are available to catalogers and end users. Figures 3 and 4 show the cataloging interface and the interface for end users to navigate the digitized objects, respectively.

It is important to note that in order for the repository to display the 3D models, 3DHOP[22] had to be adapted to Tainacan via Wordpress by installing and activating a plugin. The plugin developed[23] is an ongoing study within the scope of this project and functions as an experiment aimed at exploring ways to connect 3D objects with digital repository management functions. Future research could develop additional functionalities and integrate other file formats in the environment.

Analysis category		Dublin Core	VRA Core	LIDO	Media Ontology	M3O	COMM	M3 Multimedia	Linked.art	CRMdig
Representation language		RDF/XML RDF *Schema*	XML *Schema*	XML *Schema* RDF/XML	OWL DL	OWL	OWL DL	OWL DL	RDF *Schema* JSON-LD	RDF *Schema*
Code clarity		Clear nomenclatures for the classes and properties of the schema.	Clear nomenclatures for the elements and attributes of the schema.	Clear nomenclatures for the elements and attributes of the schema.	Clear nomenclatures for the concepts.	Entities organized by ontology content design patterns.	Complex structure derived from an upper ontology.	Clear nomenclatures for the concepts.	Well-organized structural elements via JSON syntax.	Clear nomenclatures for the classes and properties of the schema.
Knowledge extraction suitability		Simplified and comprehensive schema consisting of 15 metadata elements that can be organized into classes.	Schema with 19 metadata elements to describe characteristics of the collection, artwork and image.	Simple terminology displayed in tree format, including 9 areas covered by the standard.	Simple taxonomy representing a set of media classes and specific non-media elements.	Simple diagrams containing modularization for multimedia and provenance standards.	Ontology partitioned into modules based on multimedia standards.	Ontology partitioned into modules representing multimedia, multidomain and multilingual characteristics.	Class diagrams representing the taxonomic relationship between the ontologies involved in the model.	Class diagrams representing the taxonomic relationship between the ontologies involved in the model.
Naming convention suitability		W3C metadata standard.	Internationally recognized metadata standard approved by METS.	CDWA Lite; Museumdat; SPECTRUM XML schema; CIDOC CRM.	W3C metadata standard.	Terminology adopted by the project team; DUL.	MPEG-7; DOLCE.	W3C metadata standard. MPEG-7.	CIDOC CRM	CIDOC CRM
Nature of the multimedia metadata and the artwork itself	Content independent	Creation, production, classification, information and use of the digital resource.	Creation, production, classification, information and use of the resource (Work, Image and Collection).	Creation, production, classification, information and use of digital resource.	Creation, production, classification, information and use of media.	Information on media.	Information on media.	Creation, production, classification, information and use of media.	Creation, production and classification, of the original and digital resource.	Creation, production and classification, of the original and digital resource.
	Content dependent	Absent	Absent	Absent	Absent	Absent	Visual elements	Visual elements	Absent	Digital measurement elements.
	Descriptive Content	Possibility of linking with other semantic vocabularies to describe resource content.	Absent	Describing events that aggregate real-world entities.	Annotation for media segments.	Multimedia standards for semantic content annotation, including segmentation and collection.	Multimedia standards for semantic content annotation, including segmentation.	Annotation for audiovisual segmentation, semantic and personalized content.	Model centered on abstract entities to represent a cultural object.	Annotations for the data acquisition event and digital object itself, as a whole or in parts.
Semantic features	Axioms	Absent	Absent	Absent	Present [a]	Present	Present	Present	Present	Present
	Annotations	Present	Present	Present	Present	Present	Present	Present [b]	Present	Present

[a] presence of axioms with simple logical statements.[b] annotations present, but without consistent conceptual statements.

Table 2. Vocabulary analysis summary.

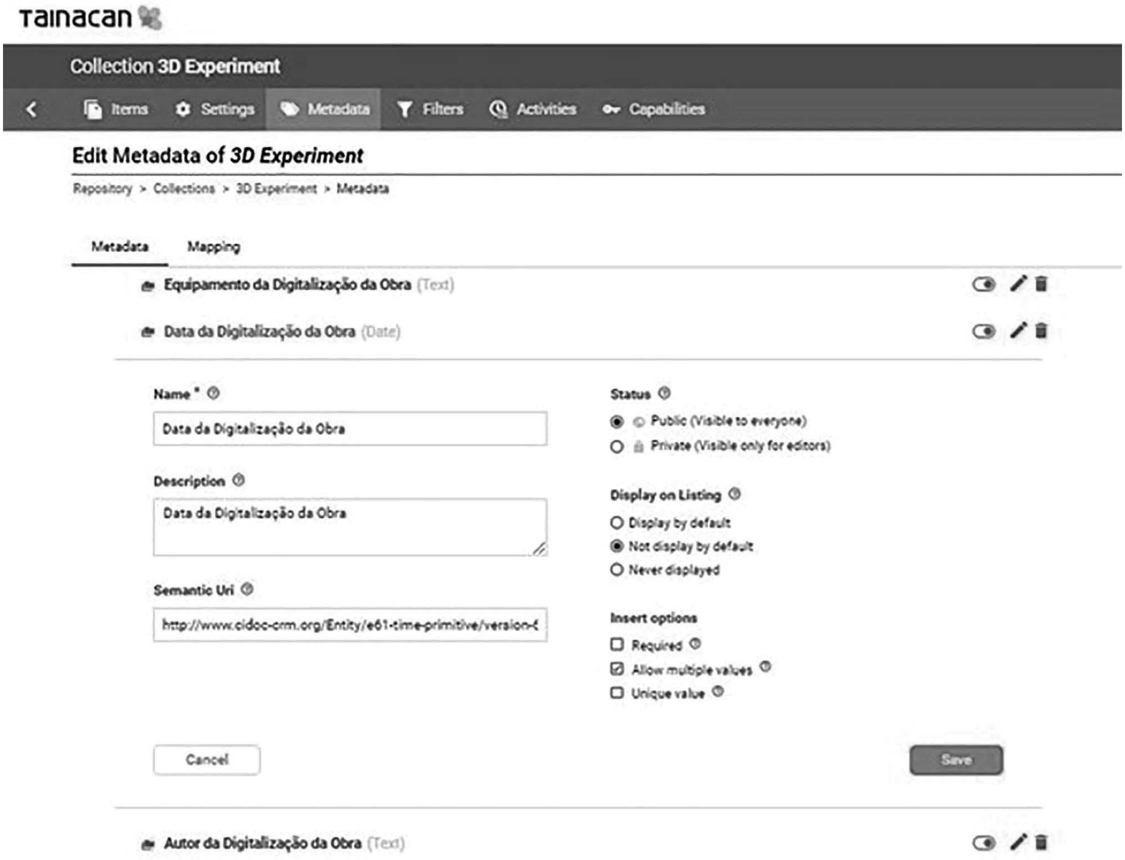

Figure 2. Example of metadata configuration with semantic identifier attribution (in Portuguese).

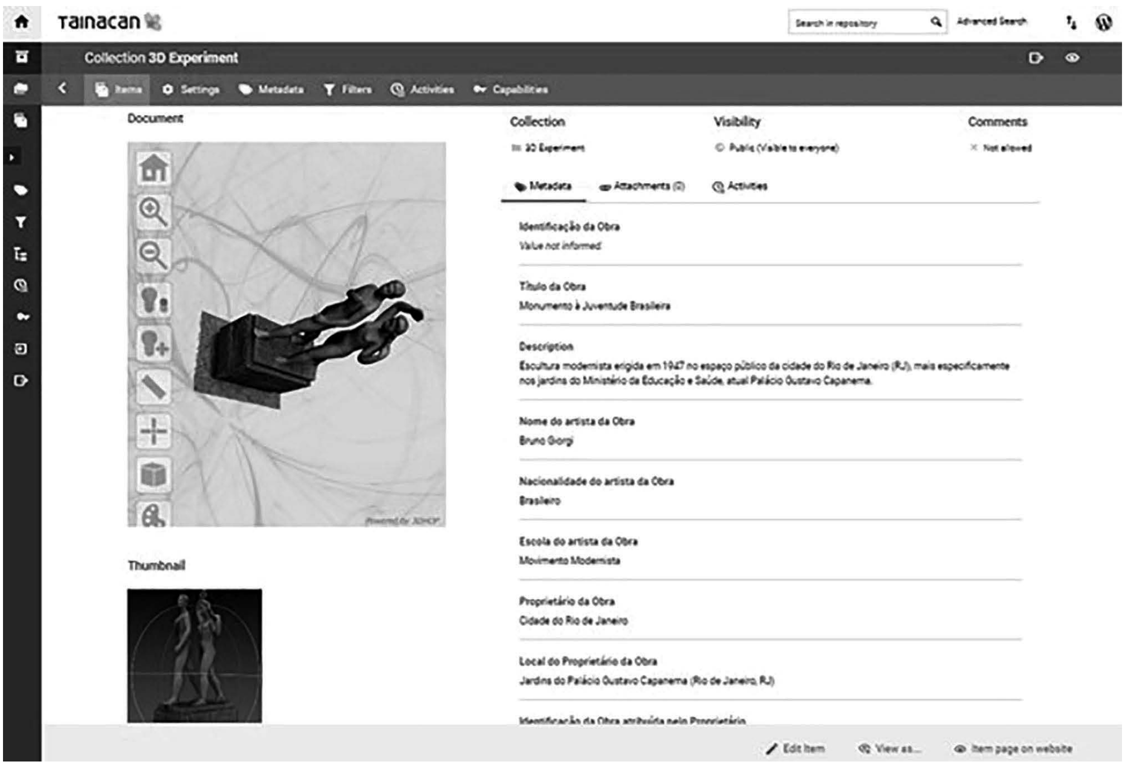

Figure 3. Administrative Interface for cataloging digital objects (in Portuguese).

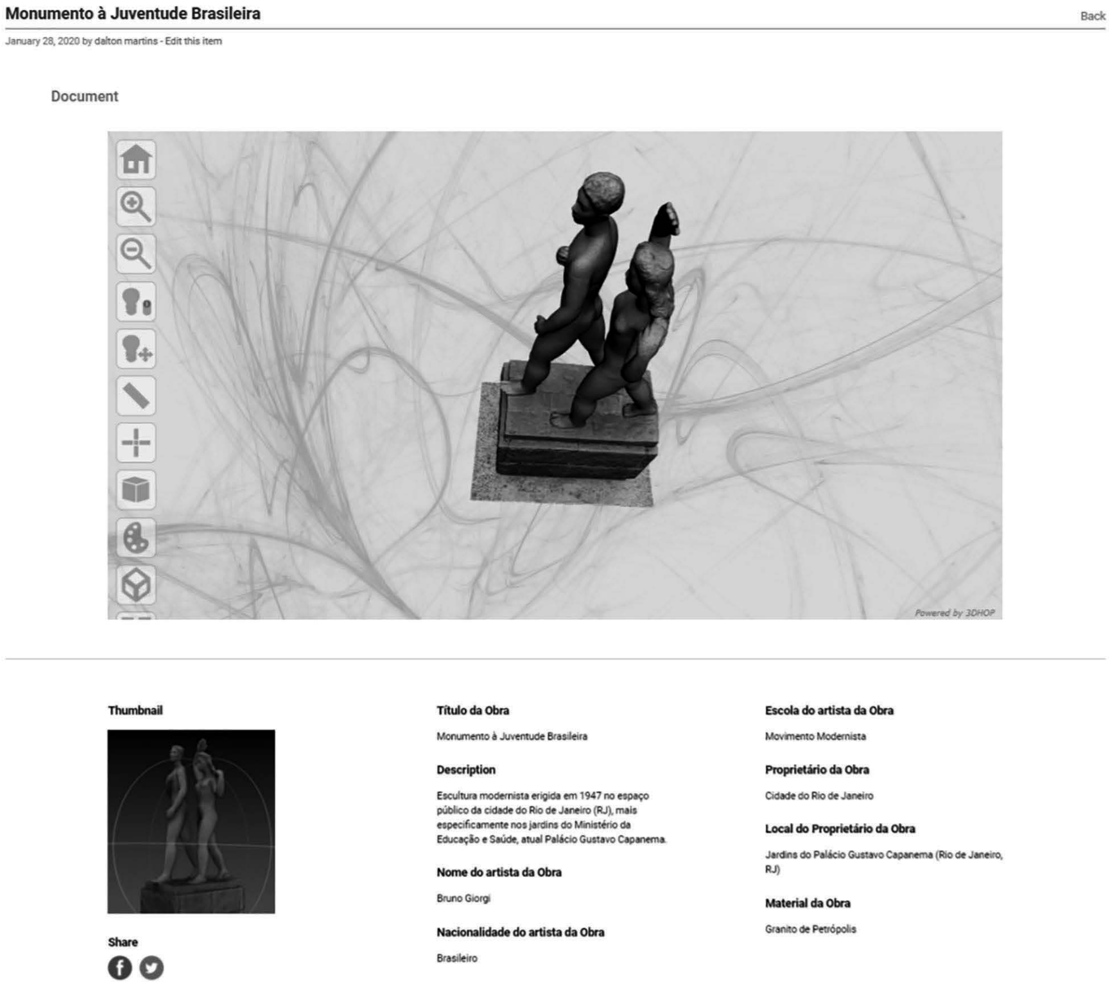

Figure 4. Visualization of a digital object and its metadata for end users (in Portuguese).

Once the objects are available on the platform, cataloged and semantically defined, Tainacan provides an Application Programming Interface (API) that enables all the metadata to be accessed in JSON format. Since the API exports the entire configuration model for each metadata element used in the cataloging form, the URIs of each element can be identified and integrated into an ontology explaining the entities and relationships to compile a representation graph of each digital object. Figure 5 shows the result of exporting documentation from the Tainacan API in JSON format to the Digitization Date metadata. This makes it possible to represent the complete object using the CRMdig ontology and construct its graph in an RDF model.

This experiment enabled a low technical complexity application and easy to implement model, ensuring a minimally technical software and creation of a highly expressive ontology for the cultural heritage universe. The software guarantees unique semantic identification of each metadata element and exports the data in an interoperable format that can be easily accessed and processed by a machine. Data could easily be collected from different repositories configured in this way, ensuring their interoperability and the possibility of building aggregated search engines and analysing information.

5.0 Conclusions and future work

This exploratory experimental study analysed and compared semantic models to create a 3D repository for modernist artworks in the city of Rio de Janeiro, Brazil. Based on a conceptual discussion aimed at investigating different analytical models, vocabularies, metadata standards and ontologies, this study created a comparability reference to present patterns that allowed representation on three analysis levels: content independent, content dependent and descriptive metadata. The proposed analytical model was constructed using Neon Methodology, which was robust and efficient at explaining different dimensions and variabilities in the

```
"data-da-digitalizacao-da-obra": {
  "name": "Data da Digitalização da Obra",
  "id": 114,
  "date_i18n": "October 20, 2019",
  "value": [
    "2019-10-20"
  ],
  "value_as_html": "October 20, 2019",
  "value_as_string": "October 20, 2019",
  "semantic_uri": "http://www.cidoc-crm.org/Entity/e61-time-primitive/version-6.2.1",
  "multiple": "yes",
  "mapping": [

  ]
},
```

Figure 5. JSON for metadata exportation.

analysis of the vocabularies identified in the literature and bibliographical review.

Analysis and comparability of semantic models provides the necessary conditions to select and reuse appropriate knowledge resources to represent a comprehensive structure capable of modeling metadata to semantically organize them into digital repositories. As discussed, knowledge resources such as COMM and M3O can supply generic concepts from multimedia design patterns along with metadata classes (independent content, dependent content and descriptive) covered by ontologies (M3 Multimedia and Media Ontology) that provide clear separation of interests in relation to media, including content semantics, knowledge related to information resource management, structural aspects of content and characteristics of the documentary reality of the multimedia type.

In the specific case of the experiment presented, we opted initially to work with the LIDO standard as an easily accessible metadata model already known by the community of cultural heritage experts that enables connection and integration (due to its representation in RDF) with multimedia ontologies from specific domains, particularly the CIDOC CRM model and its extensions, including CRMdig for provenance scenarios and Linked.art for smoother insertion (in terms of usability) in the LOD movement.

In this respect, the proposal of documenting open spaces cultural heritage assets based on well-established metadata standards and ontologies recommended by communities and regulatory bodies such as W3C and ISO is a unique and strategic approach in the search for intelligent solutions to describe machine-processable semantic-based multimedia content. Thus, the aspects of the first set of digitized replicas in the project (RWO, 3DR and 3DO) were described using the recommended vocabularies.

The delivery and consumption of 3D digital replicas through a repository envision the need for semantic integration of these multimedia resources in a network, allowing sharing and linking between several collections on the web. In this respect, the 3D Digital Heritage: Modern Rio project aims to provide the community with a digital repository that allows users free access to collections of modernist cultural heritage monuments and sculptures in Rio de Janeiro.

The free software (Tainacan) used for the repository was easy to configure and implement, since it only requires defining the metadata model and attributing a semantic identifier to each metadata element. After configuration, the digitized 3D objects were cataloged. The data were exported in JSON format via an API, aggregated and configured to represent and RDF file, which could then interoperate with other repositories, demonstrating its viability for semantic interoperability.

We believe that the proposed solution represents a highly expressive model for 3D cultural heritage objects and is an accessible, easily implemented technology based on a widely known metadata model, with a large internet user base (WordPress). As such, we hope to contribute to the viability of implementing semantic digital repository solutions accessible even for small or low-tech institutions, where there is still a lag in adopting these information representation and semantic interoperability models.

In regard to future studies, this is an ongoing project with room for additional avenues. Using the proposed methodology, historians and curators worked on the informational context and links to other documentation bases in order to enhance the narrative of modern and contemporary heritage in Rio de Janeiro. The semantic repository on Tainacan will continue to be developed and improved through a series of tests related to metadata handling and ontologies aimed at semantic annotation of digital 3D replicas. Additionally, the conceptual model initially proposed based on LIDO and CRMdig will also be improved in order to integrate information about the digitized replicas in their original aspects in terms of provenance, media and 3D content, with the various collections of the cultural institutions available in the LOD semantic network. Based on this integration, different SPARQL queries with useful inference mechanisms will be implemented and tested in a triplestore to obtain more conclusive results about the semantic repository created in Tainacan.

Notes

1. MUSEUM P.: 3d Petrie museum. https://www.ucl.ac.uk/3dpetriemuseum/3dobjects
2. STARC: Starc. http://public.cyi.ac.cy/starcRepo/explore/objects
3. 3D-ICONS: 3d-icons. http://3dicons-project.eu/
4. FGV: Center for Research and Documentation of Contemporary Brazilian History. https://cpdoc.fgv.br/sobre
5. FGV: School of Applied Mathmatics. https://emap.fgv.br/
6. In the context of the web, a resource is any artifact identifiable by a single identifier (e.g. a URI – Uniform Resource Identifier), such as digital documents expressed in different media.
7. https://wiki.dbpedia.org/
8. Amon Carter Museum of American Art; Archives of American Art, Smithsonian Institution; Autry Museum of the American West; Colby College Museum of Art; Crystal Bridges Museum of American Art; Dallas Museum of Art (DMA); Indianapolis Museum of Art (IMA); Thomas Gilcrease Institute of American History and Art; National Portrait Gallery, Smithsonian Institution; National Museum of Wildlife Art; Princeton University Art Museum; Smithsonian American Art Museum (SAAM); The Walters Art Museum and Yale Center for British Art.
9. https://linked.art/index.html
10. https://json-ld.org/
11. https://www.getty.edu/research/tools/vocabularies/
12. Inventory of Monuments in Rio de Janeiro. http://inventariodosmonumentosrj.com.br/
13. http://www.dublincore.org/specifications/dublin-core/dcq-rdf-xml/
14. https://www.loc.gov/standards/vracore/schemas.html
15. http://www.loc.gov/standards/mets/
16. http://vraweb.org/vra-core-rdf-ontology-available-for-review/
17. http://www.lido-schema.org
18. http://mpeg.chiariglione.org/standards/mpeg-7
19. http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/completedprojects/66-buscamedia/index.html
20. http://www.oeg-upm.net/
21. https://wordpress.org/plugins/tainacan/
22. Open source framework for the creation of interactive Web presentations of high-resolution 3D models. Conf. :http://vcg.isti.cnr.it/3dhop
23. https://github.com/tainacan/wordpress-3dhop

References

Abbas, June. 2010. *Structures for Organizing Knowledge: Exploring Taxonomies, Ontologies, and other Schemas*. New York: Neal-Schuman.

Alexiev, Vladimir. 2018. "Museum Linked Open Data: Ontologies, Datasets, Projects." In *Digital Presentation and Preservation of Cultural and Scientific Heritage. Conference Proceedings* 8, Sofia, Bulgaria: Institute of Mathematics and Informatics – BAS, 19–50. http://dipp.math.bas.bg/images/2018/019-050_32_11-iDiPP2018-34.pdf

Almeida, Maurício Barcellos. 2013. "Revisiting Ontologies: a Necessary Clarification." *Journal of the American Society for Information Science and Technology* 64: 1682-93.

Arp, Robert, Barry Smith, and Andrew D. Spear. *Building ontologies with basic formal ontology*. Cambridge, MA: MIT Press, 2015.

Benjamins, V. Richard, Mark Radoff, Mike Davis, Mark Greaves, Rose Lockwood and Jesús Contreras. 2011. "Semantic Technology Adoption: A Business Perspective." In: *Handbook of Semantic Web Technologies*, edited by John Domingue, Dieter Fensel and James A, Hendler. Berlin Heidelberg: Springer, 619-57.

Bizer, Christian, Tom Heath and Tim Berners-Lee. 2009. "Linked Data - The Story So Far." *International Journal on Semantic Web and information systems* 5:1–22. doi: 10.4018/jswis.2009081901

Borgo, Stefano and Claudio Masolo. 2009. "Foundational Choices in DOLCE." In *Handbook on Ontologies*, edited by S. Staab and R. Studer. Springer Berlin Heidelberg, 361-81.

Charles, Valentine, Robina Clayphan, and Antoine Isaac (Eds.). 2017. "Definition of the Europeana Data Model v5.2.8." In *Europeana: Think Culture*. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM_Definition_v5.2.8_102017.pdf

Clement, Tanya E., and Daniel Carter. 2017. "Connecting Theory and Practice in Digital Humanities Information Work." *Journal of the American Society for Information Science and Technology* 68:1385–96. doi: 10.1002/asi.23732

Doerr, Martin, Stephen Stead, Maria Theodoridou et al. 2016. "CRMdig Model for Provenance Metadata." https://www.ics.forth.gr/isl/CRMext/CRMdig/docs/CRMdig_v3.2.1.pdf.

Doerr, Martin, and Dolores Iorizzo.2008. "The Dream of a Global Knowledge Network—A New Approach." *Journal on Computing and Cultural Heritage* 1: 1-23. doi: 10.1145/1367080.1367085

Dahlberg, Ingetraut. 1993. "Knowledge Organization: its Scope and Possibilities." *Knowledge Organization*, 20: 211-22.

Echavarria, Karina Rodriguez, Karin Janzon and Jonathan Wright. 2018. "Participatory Co-creation of a Public Sculpture Incorporating 3D Digital Technologies." In *Eurographics Workshop on Graphics and Cultural Heritage*, edited by Robert Sablatnig and Michael Wimmer. The Eurographics Association. doi: 10.2312/gch.20181364

Fink, Eleanor E. 2018. "American Art Collaborative (AAC) Linked Open Data (LOD) Initiative: Overview and Recommendations for Good Practices." http://americanartcollaborative.org/wp-content/uploads/2018/03/AAC_LOD_Overview_Recommendations.pdf.

Foskett, Douglas John. 1985. "Classification and Integrative Levels." In *Theory of subject analysis*, edited by L.M. Chan, P.A. Richmond and E. Svenonius. Littleton, CO: Libraries Unlimited, 212–20.

Gilliland, Anne. J. 2016. "Setting the Stage." In *Introduction to Metadata* 3rd edition, edited by Murtha Baca. Los Angeles: Getty Research Institute. http://www.getty.edu/publications/intrometadata/setting-the-stage/

Gonizzi Barsanti, Sara and Gabriele Guidi. 2013. "3D Digitization of Museum content within the 3D-ICONS Project." In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W1: 151–6. doi: doi.org/10.5194/isprsannals-II-5-W1-151-2013

Guarino, Nicola. 1998. "Formal Ontology In Information Systems." In *Proceedings of the First International Conference (FOIS'98)*, Trento, Italy. Amsterdam: IOS Press.

Gueguen, Gretchen, Ivey Glendon, Valentine Charles, Robin Dean, Greta Bahneman, Juliet Hardesty, Nicole Lawrence, Kristen Merryman, Anna Neatrour and Hannah Stitzlein. 2017. "MetadataApplication Profile, Version 5.0." In *DPLA - Digital Public Library of America*. https://drive.google.com/file/d/1fJEWhnYy5Ch7_ef_-V48-FAViA72OieG/view

Guizzardi, Giancarlo. 2005. "Ontological Foundations for Structural Conceptual Models." PhD dissertation, Twente University of Technology.

Hendler, Jim, F. Gandon and Dean Allemang. 2020. *Semantic Web for the Working Ontologist: Effective Modeling For Linked Data, RDFS, and OWL*. 3rd edition. Association for Computing Machinery.

Hildebrand, Michiel, Jacco van Ossenbruggen, Lynda Hardman, Jan Wielemaker and Guus Schreiber. 2010. "Searching in Semantically Rich Linked Data: A Case Study in Cultural Heritage." http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=179DEB9CD4388DE9FC4074CDE7FF1BCE?doi=10.1.1.154.3789&rep=rep1&type=pdf

Hjorland, Birger. 2003. "Fundamentals of Knowledge Organization." *Knowledge Organization* 30: 87-111.

Hjorland, B. (2007). "Semantics and Knowledge Organization", *Annual Review of Information Science and Technology* 41: 367-405.

Hjørland, Birger. 2015. "Theories are Knowledge Organizing Systems (KOS)." *Knowledge Organization* 42:113–28. doi: 10.5771/ 0943-7444-2015-2-113

Hjørland, Birger. 2016. "Does the Traditional Thesaurus Have a Place in Modern Information Retrieval?" *Knowledge Organization* 43:145–59. doi: 10.5771/0943-7444-2016-3-145

Hyvönen, Eero. 2012. "Publishing and Using Cultural Heritage Linked Data on the Semantic Web." *Synthesis Lectures on the Semantic Web: Theory and Technology* 2: 1–159. doi:10.2200/S00452ED1V01Y201210WBE003

IFLA Study Group on the Functional Requirements for Bibliographic Records. 2009. "Functional Requirements for Bibliographic Records: Final Report as Amended and Corrected Through February 2009." http://www.ifla.org/VII/s13/frbr/

Knoblock, Craig A, Pedro Szekely, Eleanor Fink, Duane Degler, David Newbury, Robert Sanderson, Kate Blanch, Sara Snyder, Nilay Chheda, Nimesh Jain, Ravi Raju Krishna, Nikhila Begur Sreekanth and Yixiang Yao. 2017. "Lessons Learned in Building Linked Data for the American Art Collaborative." In *The Semantic Web – ISWC 2017*. Lecture Notes in Computer Science, Vol 10588. Cham: Springer, 263–79. https://link.springer.com/chapter/10.1007/978-3-319-68204-4_26

Koltay, Tibor. 2016. "Library and Information Science and the Digital Humanities: Perceived and Real Strengths and Weaknesses." *Journal of Documentation* 72: 781-92.

Le Boeuf, Patrick, Martin Doerr, Christian Emil Ore and Stephen Stead (Eds.) 2018. "Definition of the CIDOC Conceptual Reference Model." In *ICOM/CIDOC Documentation Standards Group and CIDOC CRM Special Interest Group, Version, 6(2)*. http://www.CIDOC CRM.org/Version/version-6.2.3-0

Lemos, Daniela Lucas da Silva and Renato Rocha Souza. 2020. "Knowledge Organization Systems for the Representation of Multimedia Resources on the Web: A Comparative Analysis." *Knowledge Organization* 47:300–19. doi:10.5771/0943-7444-2020-4-300

Machado, Luis Miguel Oliveira, Renato Rocha Souza and Maria da Graça Simões. "Semantic Web or Web of Data? A Diachronic Study (1999 To 2017) of the Publications of Tim Berners-Lee and the World Wide Web Consortium." *Journal of the American Society for Information Science and Technology* 70:701–14. doi: 10.1002/ asi. 24111

Martins, Dalton Lopes, José Murilo Costa Carvalho Júnior and Leonardo German. 2018. "Projeto Tainacan: Experimentos, Aprendizados e Descobertas da Cultura Digital No Universo dos Acervos das Instituições Memoriais." In: *TIC Cultura* 59–68. https://cetic.br/media/docs/publicacoes/1/tic_cultura_2018_livro_eletronico.pdf

Martins, Dalton Lopes, Daniela Lucas da Silva Lemos, Luis Felipe Rosa de Oliveira, JoyceSiqueira, Danielle do Carmo and Vinicius Nunes Medeiros. 2022. "Information Organization and Representation in Digital Cultural Heritage in Brazil: Systematic Mapping of Information Infrastructure in Digital Collections for Data Science Applications." *Journal of the Association for Information Science and Technology* 73:1–20. https://doi.org/10.1002/asi.24650

Medeiros e Sá, Asla, Adolfo Ibañez Bartolome, Karina Rodriguez-Echavarria, Ricado Marroquim and Vivian Luiz Fonseca. 2019. "Accessible Digitisation and Visualisation of Open Cultural Heritage Assets." In *Eurographics Workshop on Graphics and Cultural Heritage*, edited by Selma Rizvic and Karina Rodriguez Echavarria. The Eurographics Association, 59-68. doi: 10.2312/gch.2019 1349.

Neely, Liz, Anne Luther and Chad Weinard. 2019. "Cultural Collections as Data: Aiming for Digital Data Literacy and Tool Development." In *MW19, the 23rd Annual MuseWeb Conference, in Boston, April 2-6, 2019.* https://mw19.mwconf.org/paper/cultural-collections-as-data-aiming-for-digital-data-literacy-and-tool-development/

Newbury, David. "LOUD: Linked Open Usable Data and linked.art." 2018. https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2021/03/14_10_david_newbury_presentation.pdf

NISO (National Information Standards Organization). 2005. *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. ANSI/NISO Z39.19-2005.* Baltimore: NISO.

Pattuelli, Maria Cristina. 2011. "Modeling a Domain Ontology for Cultural Heritage Resources: A User-Centered Approach." *Journal of the American Society for Information Science and Technology* 62:314–42. doi: 10.1002/asi.21453

Pitzalis, Denis, Franco Niccolucci and Matthieu Cord. 2011. "Using LIDO to Handle 3D Cultural Heritage Documentation Data Provenance." In *9th International Workshop on Content-Based Multimedia Indexing (CBMI).* IEEE, 37–42. doi: 10.1109/CBMI.2011.5972 517

Poole, Alex H. 2016. "The Conceptual Ecology of Digital Humanities." *Journal of Documentation* 73:91–122. doi: 10.1108/JD- 05- 2016- 0065

Potenziani, Marco, Marco Callieri, Matteo Dellepiane, Massimiliano Corsini, Federico Ponchio and Roberto Scopigno. 2015. "3DHOP: 3D Heritage Online Presenter." In: *Computers & Graphics* 52, 129–41. doi: https://doi.org/10.1016/j.cag.2015.07.001

Potenziani, Marco, Marco Callieri, Matteo Dellepiane and Roberto Scopigno. 2018. "Publishing and Consuming 3D Content on the Web: A Survey." *Foundations and Trends® in Computer Graphics, and Vision* 10(4): 244-333. doi: 10.1561/0600000083

Rademaker, Alexandre, Dário Augusto Borges Oliveira, Valeria de Paiva, Suemi Higuchi, Asla Medeiros e Sá and Moacyr Alvim. 2015. "A Linked Open Data Architecture for the Historical Archives of the Getulio Vargas Foundation." *International Journal on Digital Libraries* 15, nos.2-4: 153–67. doi: 10.1007/s00799-015-0147-1.

Schandl, Bernhard, Bernhard Haslhofer, Tobias Bürger, Andreas Langegger and Wolfgang Halb. 2012. "Linked Data and Multimedia: The State of Affairs." *Multimedia Tools and Applications* 59: 523-56.

Scopigno, Roberto, Marco Callieri, Matteo Dellepiane, Federico Ponchio and Marco Potenziani. 2017. "Delivering and Using 3D Models on the Web: Are We Ready?" *Virtual Archaeology Review* 8(17): 1–9. doi: 10.4995/var.2017.6405

Shadbolt, Nigel, Tim Berners-Lee and Wendy Hall. 2006. "The Semantic Web Revisited." *IEEE Intelligent Systems* 21: 96-101.

Silva, Daniela Lucas da. 2014. "Ontologias para Representação de Documentos Multimídia: Análise e Modelagem." PhD dissertation. Universidade Federal de Minas Gerais.

Suárez-Figueroa, Mari Carmen, Asunción Gómez-Pérez and Mariano Fernández-López. 2012. "The NeOn Methodology for Ontology Engineering." In *Ontology Engineering in a Networked World*, edited by Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta and Aldo Gangemi. Berlin: Springer, 9- 34.

Zeng, Marcia Lei, and Jian Qin. 2016. *Metadata*. 2nd edition. Atlanta, GA: ALA Neal-Schuman.

Zeng, Marcia Lei. 2019. "Interoperability." *Knowledge Organization* 46: 122-46.