



Sharing the knowledge: exploring cultural heritage through an ontology-based platform

Mario Casillo¹ · Massimo De Santo¹ · Rosalba Mosca¹ · Domenico Santaniello¹

Received: 29 December 2022 / Accepted: 30 May 2023 / Published online: 30 June 2023
© The Author(s) 2023

Abstract

The exponential growth of data has made ontology modeling an up-and-coming research area for knowledge representation. One of the domains of interest is represented by Cultural Heritage, in which modeling of environments, such as archaeological parks, could lead to their protection and enhancement. In this scenario, the valuable dataset on the Naples Urban Archaeological Park (PAUN) is of particular interest. Due to its peculiarity, the database could benefit from innovative techniques for retrieving the information, representing a large part of the information contained within the DatabencArt platform, which collects data on Campania's cultural heritage. This paper aims to introduce an ontology-based approach to improve data retrieval, which could help expert users in the field (archaeologists, art historians, geologists, etc.). Scholars need to be able to easily compare the information. To this end, a semantic search is able to transform a vast amount of data into linked concepts. We mapped all concepts according to the ArCo ontology and equipped the DatabencArt platform with a semantic query layer, making the information accessible through ArCo ontology's properties. To evaluate the proposed approach, we designed an experimental campaign involving expert users to assess the results obtained in the system's queries through Sparql queries. The obtained results are promising.

Keywords Ontology · Cultural heritage · Experiential learning · Knowledge management

1 Introduction

Information Retrieval techniques are often limited to keyword searches. This traditional information retrieval model is the basis of most (primarily archival) information systems today (Rasmussen 1999). This is also the primary mode of operation of search engines that allows finding Web information. Many theories and models have keyword search as their basis: one of the most frequently used and cited information search paradigms, Ben Shneiderman's Visual Information-Seeking Mantra, uses the traditional Information Retrieval model as its starting point (Shneiderman 1996). However, this model has a few limitations.

The scientific community is sharply divided between scholars in the 'hard' sciences, who prefer searching through query formulation, and scholars in the humanities and social sciences, who prefer more informal information retrieval

methods (Cole 2000; Bates 2005; Hau et al. 2005). Moreover, this model often provides a long, disorganized list of search results that do not directly address users' information needs (Gendt et al. 2006). Improving information search in the cultural heritage sector is an active area for Semantic Web researchers (2007; Aroyo et al. 2007).

This paper presents our work to improve the Information Retrieval technique from the cultural heritage database in Campania (DatabencArt). The aim was to make the data contained in the platform usable in a way that would be useful to expert users in the field (archaeologists, art historians, geologists, etc.). Scholars need to be able to compare the information easily. To this end, a semantic search transforms an island of data into linked concepts.

The primary source of data we have referred to concerns the domain of the Urban Archaeological Park of Naples, stored in our DatabencArt database in the form of records processed on the ICCD (Central Institute for Catalogue and Documentation) standard. To achieve our aim, we mapped all the fields of the cards onto the nodes and properties of the ArCo ontology. The DatabencArt platform was equipped

✉ Rosalba Mosca
rmosca@unisa.it

¹ DIIN University of Salerno, Fisciano, SA, Italy

with a semantic query layer so that the information assets could be queried via the ArCo ontology.

To evaluate the system, we designed an experimental campaign involving expert users able to define some search criteria suggested and compared them with the results obtained in the system's queries through Sparql queries.

The article is organized as follows. The following section presents the Naples Urban Archaeological Park (PAUN) and the DatabencArt platform in which the Park's area of intervention data have been collected. Then, a literature review based on ontologies to improve the semantic consultation of data is reported. The following section presents the proposed system to improve the technique for retrieving information on the cultural heritage of Campania. To describe the system, we analyze the architecture of the DatabencArt platform and the Semantic Module (SeM) developed. Next, we present the methodology of our work, which is based on semantic queries built through constant dialogue with experts in the field. This section also presents an overview of the concrete mapping between the ICCD format records contained in DatabencArc and ArCo, the reference ontology chosen for data integration. The following section reports an experimental campaign to test the effectiveness of responding to queries formulated by experts. Finally, the paper ends with a conclusion section, which discusses the results obtained and further studies.

2 Background

In recent years, Information Systems have been implemented with data flows and substantial datasets on the heritage, mainly archaeological, from excavations for underground mobility (Colace et al. 2015). In the context of Naples, the realization of the underground offered the opportunity to implement an experimentation process to develop a valorization and protection model with the design of the Urban Archaeological Park. The PAUN project (<https://paun.databenc.it>, 2023) financed with European funds by the Campania Region, is based on the assumption that without knowledge, there is no lasting valorization. Without protection, there is no knowledge or valorization to pass on to future generations.

The design of the Municipal Station led to the planning by the Superintendency of an urban park that represents an almost unique example at the national and international levels.

As part of this highly advanced strategy, the PAUN project is part of a collaboration with the Municipal Administration and the Naples Metropolitan Company and a specific agreement with the *Soprintendenza Archeologia, Belle Arti e Paesaggio* for the City of Naples.

The intervention area corresponds to the north-western sector of the interchange between *Castel Nuovo*, the gardens of *Piazza Municipio*, the underlying archaeological site in the station of metro lines 1 and 6, and the church of the *Incoronata*.

The context thus defined offers a sample for an experiment that privileges diachrony.

The aim is to enhance and account for the multi-layered life of the site and the transformations of an crucial sector of its coastal strip, marked by the port from the archaic age to the present day.

The survey data led to the creation of the Integrated Knowledge System through acquisition and management in the DATABENCART platform. These data are related to valuable results: approximately 40,000 surveys, orthophotos and photos, point clouds, and 3D models; the creation of maps of archaeological/monumental potential and archaeological/environmental risk/disturbance in order to develop predictive safeguard models (WEBGIS).

Data from different specializations and investigation methodologies contribute to the construction of the Integrated Archaeological, Historical-Monumental, and Landscape Knowledge System, in particular, the archaeological data acquired with the new excavations for the construction of the metro lines 1 and 6, available thanks to the collaboration with the ABAP Superintendency for the City of Naples. All information, including that from pre-existing databases, is normalized based on the cataloging and formal standards provided by SITAN and SIGEC-WEB and is managed by a platform for the collection, cataloging, and hierarchical structuring of data.

3 Related works

As the literature suggests, ontologies and ontology-based metadata have the potential to significantly improve information systems, including libraries and digital data archives or educational applications (Moraitou et al. 2019; Moraitou et al. 2023).

This work starts with studying tools that facilitate using Campania's digital cultural heritage resources by researchers and experts.

In the field of databases and information integration, there is a range of research to facilitate interoperability between different systems (Giallonardo et al. 2019). This research explores techniques for consulting databases through semantic queries.

In this field, ontology is fundamental because it semantically organizes data.

The most cited definition of ontology is “a formal and explicit specification of a shared conceptualization” (Gruber 1993). A conceptualization refers to an abstract, simplified

view of a portion of the world one wants to represent to achieve a previously established purpose. An explicit specification means that the concepts and relations of the abstract model are given explicit names and definitions. Formal means that the specification of meaning is encoded in a language based on logic. Formality is a critical way to eliminate ambiguity in natural language; it also opens the door for automatic inference to derive new information from the specification of meaning. Shared means that the main purpose of an ontology is generally to be used and reused in different applications and communities.

A detailed discussion on the uses of ontologies and the challenges of the scientific community for semantic integration is in the article by Uschold, Michael, and Michael Gruninger (Uschold and Gruninger 2004). The authors identify the Holy Grail of achieving fully automatic semantic interoperability. They suggest the use of common standards, such as a shared top-level ontology, as the best pragmatic approach. They also call for improved mapping techniques by investigating interactive methods involving humans.

Many solutions have been proposed to achieve efficient interoperability between heterogeneous information systems. The article by Cullot, Nadine, Raji Ghawi, and Kokou Yétongnon (Cullot et al. 2007) focuses on mappings between databases and ontology. To this end, a tool called db2OWL, was developed to create ontologies from a relational database. Each database component (table, column, constraint) is converted into a corresponding ontology component (class, property, relation).

Another method was proposed by Hazber et al. (Hazber et al. 2015) and consists of two main steps: constructing ontologies from relational database data and automatically generating ontology instances from database data.

A fascinating work is the one by Sir et al. (Sir et al. 2015), whose aim is to clarify the differences between ontologies and databases. The article describes step by step, the parts where the differences occur. However, some similarities show that ontologies and databases are not entirely different. Based on these aspects, the article presents various approaches to transforming a database into an ontology. The conclusion summarises and highlights the most important similarities and differences.

Another interesting paper is proposed by Capuano et al. (2015), which categorizes how semantic technologies can support information retrieval in interactive visualizations; Chang et al. (2020) focus on the mapping and automatic construction of ontologies from relational databases.

This work examined the relationships between our NoSQL database schema DatabencArt (in ICCD card format) and the formal ontology ArCo (Carriero et al. 2019).

ArCO (Architecture of Knowledge in Italian) is a network of ontologies for structuring highly expressive cultural heritage knowledge. The data in ArCo are structured based on

ontological models that mirror the structural analyticity of the ICCD ministerial records used to describe cultural heritage. Precisely because of the consistency with the structural analyticity of the ICCD forms, we chose ArCo's ontology to perform a mapping exercise that would allow us to interrogate our data through a semantic level.

This article is not intended to provide a comprehensive review of the state of the art of using ontologies for semantic integration. For this purpose, we refer the reader to the in-depth study by Kalfoglou and Schorlemmer (2003) and the latest new contributions in the literature on this topic (Ranjgar et al. 2022; Krabina 2023).

Instead, in this article, we want to emphasize the importance for specialists in the field of Cultural Heritage (CH) to compare their results with the standards that encode them.

In fact, Cultural Heritage encompasses a wide range of disciplines and traditions that relate to the human past and needs better data integration.

4 The proposed approach

This article aims to present a system to improve the technique for retrieving information on the cultural heritage of Campania through the use of Ontologies. Our approach, which aims to improve data retrieval, is based on how an expert searches for information and, above all, on their research objectives. The experts who have supported us in this work belong to different fields; since the type of data cataloged in DatabencArt is the result of the cataloging of artifacts involving numerous sciences (archaeology, geology, art history, chemistry, etc.).

Digitization activities have made it possible to document material relating to the Urban Archaeological Park of Naples that was mainly found during excavations for the metro construction. The data were entered into the platform upon acquisition to facilitate consultation, study, and research. These data, therefore, represent an essential source, fundamental for processing what was recovered in the field research. However, mere storage of information filed in ICCD format does not lend itself well to data analysis and interpretation. On the other hand, the conversion and management of data through a semantic layer allow the information to be related and uncover novel links and divergent exploratory paths from the initial hypotheses.

In this regard, the following sections will present the architecture of the DatabencArt platform, developed by the DATABENC Consortium, which aims to use information technologies for conserving, valorizing, and sustainable use of Cultural Heritage. Currently, the search allowed on the platform is syntactic, i.e., a keyword search. An expert can therefore find the file of the individual artifact and consult it. To relate it to the others, he will have to read all the cards he

is interested in and do some linking work. In this scenario, examining each card can be tiring and involves long consultation times, so an additional module, capable of interacting at several levels of the DatabencArt platform, will be presented that allows semantic type searching. This type of search, in which data are organized using an ontology, is undoubtedly helpful for an expert user, making consultation strategies more efficient. A fundamental characteristic of semantic-type searches is that they can lead to the discovery of new relationships between data. A user, for instance, may find an artifact with several other related ones. This allows them to gain further context concerning the search criteria. We can say that a search through semantic queries allows for developed contextual knowledge. The development methodology of our work was based on an ongoing dialogue with experts in the field, which highlighted essential guidelines to improve the consultation and study of the Urban Archaeological Park of Naples. The fruit of this interaction enabled domain mapping through ontologies, which will be explored in more detail below.

4.1 Databencart platform architecture

The DatabencArt platform was developed by the DATA-BENC Consortium, which aims to integrate new information technologies for conserving, valorizing and sustainable use of Cultural Heritage (Cornevilli et al. 2020). The choices made in the development of the platform, and the synthesis of several comparisons with numerous stakeholders involved, relate to the use of standards and the need for flexibility imposed by the broad complexity of the Cultural Heritage domain. In particular, the platform uses the ICCD v3.0 cataloguing standard, with the availability of managing multilingual and multiprofile descriptive fields and attributes. The platform uses NOSQL technologies and APIs for the exchange of information with the outside world. In addition, it is equipped with a smart Front-End, suitable for use on different types of devices, and a Back-End that guarantees multi-tenancy to facilitate third-party cataloguing efforts. The platform is fully integrated with IoT devices that can be integrated for the purpose of monitoring or safeguarding Cultural Heritage.

Figure 1 shows the different layers that make up the architecture of the DatabencArt platform. In the lower layer (Data Layer) all content is stored, such as information on users, cultural heritage, multimedia content, and links to Linked Open Data (LOD). This is managed by an RDBMS (Relation Database Management System) repository, a CMS (Content Management System), a NoSQL database (ElasticSearch) dedicated to the management of all content, and an Object Storage for the storage of multimedia content.

The Data Access Manager, located above the Data Layer, is dedicated to access management. This layer allows the

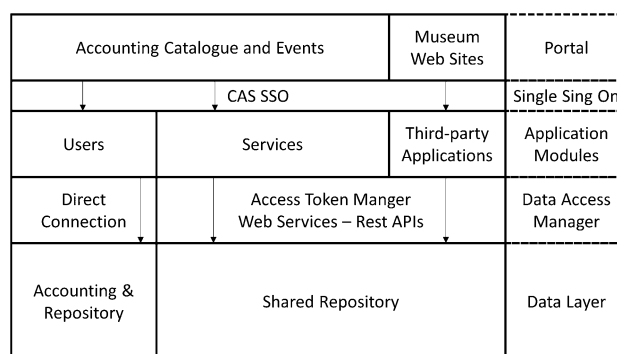


Fig. 1 DatabencArt platform levels

different applications and services of the platform to access the data they need, depending on the access scope, via REST API services. The presence of an Access Token Manager guarantees protection against unauthorized access.

The Application Layer includes modules for user management, platform management (publication, data entry, search, etc.), and external modules usable via REST API services. The Sign-On layer provides single sign-on functionality through the CAS (Central Authentication Service) protocol. Finally, the last layer is represented by the GUI layer, in which the web portal for accessing and using services within the platform or the websites of the museums associated with the platform is located.

In order to provide the platform with the possibility of semantic queries, a further module called Semantic Module (SeM) was introduced, capable of interacting with the platform at the Data Layer level. This module will be presented in detail in the following section.

4.2 Semantic module (SeM)

The Semantic Module (SeM) development was pursued to make the information heritage present in the DatabencArt platform searchable in a semantic way. Furthermore, the module is designed to allow the export and import of data related to cultural heritage in Linked Open Data (LOD) logic.

In general, the objective was to equip the platform with a semantic layer to make it interoperable with other data sources, thus constituting a knowledge network, according to the Linked Open Data (LOD) logic. Concerning the LOD import/export part, the platform was equipped with a job-processing mechanism that, on a user basis, manages the various requests for the import and export of datasets, and processing logic of maximum quotas of importable datasets.

This need arose from making the platform compatible with Knowledge Graph projects representing cultural heritage, such as ArCO (<http://wit.istc.cnr.it>, arco. 2023). To this

end, the need arose to introduce a semantic database component so that the data would be represented and maintained in a format suitable for semantic queries. Therefore, the information assets enclosed within the DatabencArt platform are converted into a standard format, such as the Resource Description Framework (RDF), and organized using two types of datasets:

1. DabencArt Datasets: datasets that expose, in RDF format, one or more subsets of the Databenc information assets. These datasets are managed and maintained in an ArCO-compliant format.
2. User-Defined Datasets: datasets loaded and imported by the user. In this case, Datasets can be represented in a user-defined format provided by the user at the upload time.

Figure 2 shows the architecture of the module introduced in the DatabencART platform, which handles operations on datasets via REST API services. Semantic access to the data, on the other hand, is provided through SPARQL endpoints, which allow the execution of the relevant queries through the Semantic Engine. In particular, the query can be performed through Web Services to allow integration with third-party applications and through a visual interface (Profiled GUI) for a more direct user experience. Access to the entire semantic system, both in terms of Data-Access and in terms of operability, takes place in a controlled manner through authentication and authorization mechanisms, as is the case within the DatabencART platform, which allows data to be exposed in the form of Linked Open Data.

The inferential capacity of the presented system depends on the concepts within the platform; this capacity grows as the available data increases, which is stored in ICCD format (ICCD, <http://www.iccd.beniculturali.it>, index.phpit, 1, home. 2023). Furthermore, a delicate aspect related to the

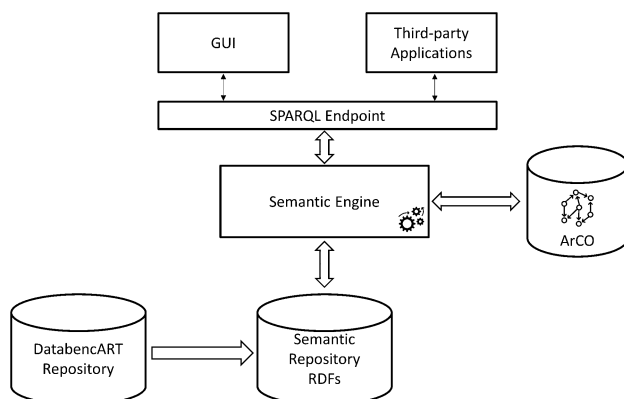


Fig. 2 Module architecture handling operations on datasets via REST API services

intersection of the concepts defined in ArCO with those in the DatabencART database lies in the design work of the ontological mapping. This aspect, which makes the integration and querying of the system possible, will be dealt with in the next section.

4.3 Method

Our approach to improving data consultation is based on how an expert searches for information and, above all, on their research objectives. The experts who have supported us in this work belong to different fields since the data cataloged in DatabencArt results from the cataloging of artifacts involving numerous sciences (archaeology, geology, art history, etc.).

Digitization activities have made it possible to document material relating to the Urban Archaeological Park of Naples that was mainly found during excavations for the metro construction. The data were entered into the platform upon acquisition to facilitate consultation, study, and research.

These data, therefore, represent an essential source, fundamental for processing what was recovered in the field research.

However, mere storage of indexed information in format does not lend itself well to data analysis and interpretation.

Conversely, the conversion of data into RDF through a semantic layer allows for the information to be related and uncover novel links and divergent exploratory paths from the initial hypotheses.

Let us examine the conceptual difference.

The search allowed on DatabencArt is syntactic, i.e., a keyword search. An expert can therefore find the record of the individual find and consult it. To relate it to the others, he will have to read all the cards he is interested in and do some linking work.

Going through each individual card can be tiring and involves long consultation times.

A semantic type of search (in which data is linked to an ontology) is undoubtedly useful for a user who already knows what he wants to find: this type of search will make the strategy more efficient.

However, a semantic search can also lead to new discoveries. A user, for instance, may find an artifact with several other related ones. This allows him to obtain further context than the one he started from. We can say that a search through semantic queries allows one to obtain developed contextual knowledge.

The methodology of our work, based on constant dialogue with experts in the field, has highlighted important guidelines to improve the consultation and study of the Urban Archaeological Park of Naples. In the section ‘[Querying data converted to RDF](#)’ we will discuss this approach in more detail.

4.4 Resources

The records were archived in DatabencArt in format.

The (<http://www.iccd.beniculturali.it>, index.phpit, 1, home. 2023) is one of the seven Central Institutes of the MiBAC, whose main objective is the creation of a centralized national catalog of the Italian cultural heritage. The Institute's activity is based on research and the development of tools, methods, and standards for the knowledge, protection, and valorization of Italy's cultural and artistic heritage.

The records in DatabencArt analysed for this work are related to Paun and are divided within the platform into two domains: Scavo Piazza Municipio and Castel Nuovo. This division was chosen by the domain experts for more efficient consultation.

Below are the types of ICCD records contained and divided by domain.

Piazza Municipio excavation:	Castel Nuovo:
MA (Archaeological Monument)	A (Architecture)
US (Stratigraphic Unit)	US (Stratigraphic Unit)
USM (Masonry Stratigraphic Unit)	USM (Masonry Stratigraphic Unit)
USR (Stratigraphic Reference Unit)	USR (Stratigraphic Reference Unit)
TMA (Table of Materials)	SAS (Stratigraphic Survey)
RA (Archaeological Find)	RA (Archaeological Find)
	OA (Work of Art)

4.5 Data to ontology mapping

The mapping process consists of defining the elements of a given data schema (source schema) using the entities

provided by another data schema (target schema) to establish a direct correspondence between the two sets of elements. Usually, mapping occurs between a closed or personal data schema and a more general and widespread standard.

We present an overview of the concrete mapping between the ICCD format records contained in DatabencArc and ArCo, the reference ontology chosen for data integration.

We began this work with a careful exploration of ArCo's ontology modules.

All record fields contained in DatabencArt were mapped onto the classes or properties in ArCo. Table 1 shows a study of the general mapping. In the table, the sign "✓" indicates the correspondence of the DatabencArt field with a concept in ArCo's ontology, while "✗" indicates no correspondence.

The mapping operation was repeated for each ICCD card type contained in DatabencArt.

Table 2 shows an example of mapping on a single card, specifically on the card of type RA belonging to the domain "Town Hall Square" and related to the movable property "Palace Giant".

4.6 Experimental campaign

In order to validate the proposed system, an experimental phase was designed, including the involvement of experienced users. The experimental phase aims to represent how, through the semantic layer, it is possible to obtain detailed answers that can support expert users.

The following paragraph presents, in detail, some possible interactions with the system, which is able to query the database through the ArCo ontology formalism by providing answers to users. In the next paragraph, taking several

Table 1 Correspondences between the modules of ArCo Ontology and records of DatabencArt

Records	Ontological fields								
	Movable/ Imovable property	Tecnichal characteristic (MTC)	Cultural domain (ATB)	State of conservation (STC)	Acces profile (ADS)	Property compiled at the time (CMP)	The class specifications of the Cultural property (OGT)	Catalogue sheets (TSK)	Legal status (TU-COND)
A	✓	✗	✓	✗	✓	✓	✓	✓	✓
CA	✓	✗	✗	✗	✓	✓	✓	✓	✓
MA	✓	✗	✗	✗	✗	✓	✓	✓	✓
MI	✓	✓	✓	✓	✗	✓	✓	✓	✓
OA	✓	✓	✓	✓	✓	✓	✓	✓	✓
RA	✓	✓	✓	✓	✓	✓	✓	✓	✓
SAS	✓	✗	✗	✗	✓	✓	✓	✓	✗
TMA	✓	✗	✗	✗	✗	✓	✓	✓	✗
US	✓	✗	✗	✓	✓	✓	✓	✓	✗
USM	✓	✗	✗	✓	✓	✓	✓	✓	✗

Table 2 Correspondences between the databencart RA type record fields and the ArCo ontology nodes and properties

DatabencArt	ArCo
Tipo oggetto/bene mobile	Arco/Cultural Entity/Bene culturale/Bene Materiale/Bene mobile
MT-Dati Tecnici	Arco/Entity/Characteristic/Caratteristica tecnica (MT/MTC)
MTC-Materia e tecnica	
AD-Accesso Dati	Denotative-description/Entity/Characteristic/Profilo di accesso (AD/ADS)
ADS-Specifiche di accesso	
CO-Conservazione	Denotative-description/Entity/Object/Concetto/Tipo/Tipo di conservazione (CO/STC/STCC)
STC	
STCC	
DT-Cronologia	Arco/Object properties/ha cronologia (DT)
DTZ	
DTZG	
CM-Compilazione	Arco/Object properties/compilato al tempo (CM/CMP/CMPD)
CMPD	
CMPN	
AU-Def Culturale	Arco/Topic/Ambito culturale (AU/ATB)
ATB	
ATBM	
ATBD	
LC-Localizzaz-geo	Arco/Object properties/ha localizzazione tipizzata nel tempo (LC)
LDC	
LDCN	
LDCU	
OG-Oggetto	Arco/Entity/Description/Specifiche del bene culturale (OG/OGT/OGTT)
OGT	
OGTN	
OGTD	

queries as a reference, the degree of satisfaction of expert users regarding the system's response is evaluated.

4.7 Querying data converted to RDF

In this section, we examine a few concrete examples that show how data conversion into RDF via a semantic layer

enables interesting semantic searches for study and research purposes.

This first example shows the interface of the semantic module that queries the data contained in DatabencArt through the ArCo ontology (Fig. 3). In the screenshot, the following Sparql query is used for the chronological attribution of assets.

```
PREFIX arco-cd: <https://w3id.org/arco/ontology/context-description/>
PREFIX arco-dd: <https://w3id.org/arco/ontology/denotative-description/>
PREFIX arco-arco: <https://w3id.org/arco/ontology/arco/>
SELECT *
WHERE{?entity rdf:type/rdfs:subClassOf* arco-arco:CulturalProperty ;
  rdfs:label ?entityLabel ;
  arco-cd:hasDating ?dating .
  ?dating arco-cd:hasDatingEvent ?event ;
  arco-cd:hasInterpretationCriterion ?criterion.
}
LIMIT 100
```

The screenshot shows the SPARQL Query Manager interface. On the left, there is a panel for 'Select Dataset' (Arco-Core) and an 'Insert Query' area containing the following SPARQL query:

```

PREFIX tiapit: <https://w3id.org/Italia/onto/TI/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT * WHERE {
?entity tiapit:atTime ?timeInterval .
?timeInterval rdfs:label ?timeLabel .
}
limit 100
    
```

Below the query is an 'Execute Query' button. On the right, the results are displayed in a table with columns 'entity', 'TimeInterval', and 'TimeLabel'.

entity	TimeInterval	TimeLabel
https://w3id.org/arco/resource/Event/1568084NV-creation-1	https://w3id.org/arco/resource/TimeInterval/secc-ix-xi	Secc. IX/XI
https://w3id.org/arco/resource/Event/1567900NV-creation-1	https://w3id.org/arco/resource/TimeInterval/secc-ix-xi	Secc. IX/XI
https://w3id.org/arco/resource/Event/1567792NV-creation-1	https://w3id.org/arco/resource/TimeInterval/xi-sec	XI sec.
https://w3id.org/arco/resource/Event/1568525NV-creation-1	https://w3id.org/arco/resource/TimeInterval/xii	XII
https://w3id.org/arco/resource/Event/1568524NV-creation-1	https://w3id.org/arco/resource/TimeInterval/xii	XII
https://w3id.org/arco/resource/ConservationIntervention/1554409NV-:onservation-intervention-1	https://w3id.org/arco/resource/TimeInterval/	-
https://w3id.org/arco/resource/ConservationIntervention/1554409NV-:onservation-intervention-1	https://w3id.org/arco/resource/TimeInterval/	..
https://w3id.org/arco/resource/ConservationIntervention/1554408NV-:onservation-intervention-1	https://w3id.org/arco/resource/TimeInterval/	-
https://w3id.org/arco/resource/ConservationIntervention/1554408NV-:onservation-intervention-1	https://w3id.org/arco/resource/TimeInterval/	..
https://w3id.org/arco/resource/ConservationIntervention/1554407NV-:onservation-intervention-1	https://w3id.org/arco/resource/TimeInterval/	-

Fig. 3 Semantic module interface that queries the data contained in DatabencArt via the ArCo ontology

As reconstructing the absolute chronology is very complex, it is important for scholars to analyze the reasons for chronological attribution in similar artifacts: the comparison allows them to understand to which period other artifacts, whose chronological attribution has not yet been completed, can be dated.

The records stored in the DatabencArt database, identified by a unique code, could be consulted individually.

The activities of an expert in the field are improved through the developed module, as it allows the semantic links of ontological structures to be exploited.

The screenshot shows the SPARQL Query Manager interface. On the left, there is a panel for 'Select Dataset' (Arco-Core) and an 'Insert Query' area containing the following SPARQL query:

```

PREFIX arco-cd: <https://w3id.org/arco/ontology/context-description/>
PREFIX arco-arco: <https://w3id.org/arco/ontology/arco/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT *
WHERE {
?arco arco-cd:hasSubject ?sub .
?sub rdfs:label ?label
FILTER(REGEX(STR(?label), "crocifisso", "i"))
} LIMIT 100
    
```

Below the query is an 'Execute Query' button. On the right, the results are displayed in a table with columns 'Sub' and 'Label'.

Sub	Label
/resource/HistoricOrArtisticProperty/1567959NV	https://w3id.org/arco/resource/Subject/b45e4aba10298d334a03432a9905d06e guest it Crocifisso, cherubini, Addolorata
/resource/HistoricOrArtisticProperty/1567936NV	https://w3id.org/arco/resource/Subject/0b688c7f07dfdd7b807eca532ddad50 guest it Crocifisso
/resource/HistoricOrArtisticProperty/1567917NV	https://w3id.org/arco/resource/Subject/0b688c7f07dfdd7b807eca532ddad50 guest it Crocifisso
/resource/HistoricOrArtisticProperty/1567933NV	https://w3id.org/arco/resource/Subject/270da9f79341ed844a880a7f118dd92f7 guest it Madonna col Bambino, Cristo Crocifisso, Santi Re Magi, Ss. Caterina d' Alessandria e Leonardo di Limoges o Antonio abate
/resource/HistoricOrArtisticProperty/1567930NV	https://w3id.org/arco/resource/Subject/2b2740c32dc96a0147131ffabe66f129 guest it Cristo Crocifisso

Fig. 4 Semantic data query using Sparql queries

A further example of semantic data querying via Sparql queries is shown in this second figure (Fig. 4). This is a thematic search for works of art. As shown in the figure, the system is able to locate assets with the subject 'crucifix'. Of course, it is possible to search for other specific subject types.

This study is useful for art historians, who often conduct subject searches on assets. A comparison with them, therefore, highlighted the need to know and analyze which assets reproduce a specific subject.

This is the sparql query used:

This is the sparql query used:

```

PREFIX arco-cd: <https://w3id.org/arco/ontology/context-description/>
PREFIX arco-arco: <https://w3id.org/arco/ontology/arco/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT *
WHERE {
?arco arco-cd:hasSubject ?sub .
?sub rdfs:label ?label
FILTER(REGEX(STR(?label), "crocifisso", "i"))
} LIMIT 100
    
```

The expert's search can be translated into numerous other queries for purposes such as reconstructing building techniques or describing wall textures, from which elements of decay can be detected for monitoring and restoration.

This search is aided by queries such as:

What is the state of conservation of cultural property X)? And what interventions for good have been proposed?

What is the construction technique of cultural property X)?

Again, comparisons with ceramics experts revealed the need for a search to determine which cultural property has yielded a type of material, e.g., black varnish. We need to know how many artifacts there are, the chronology of

attribution, the type (e.g., cups/glasses), and whether the good is imported or is local production.

This research could be translated into queries such as:

What are the goods made of black paint? What is the attribution chronology? What type are they?

The formulation of these queries allowed us to assess the system's effectiveness. In other words, the queries also serve the function of competence questions that enable us to check whether the system is able to provide adequate answers.

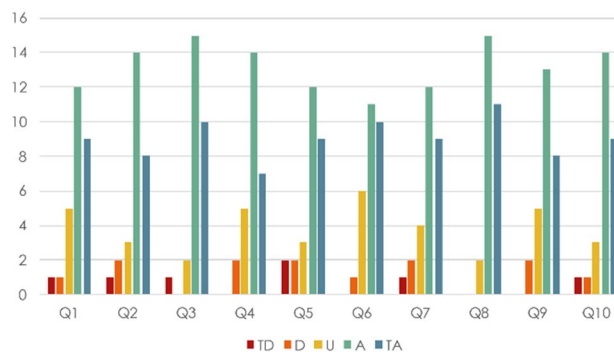


Fig. 5 Expert user validation answers

Table 3 A list of competency questions

Queries
Query 1: When did cultural property x date back to? And what motivation justifies the chronology?
Query 2: What are the assets with subject x?
Query 3: What is the state of conservation of cultural asset x? And what interventions on the property have been proposed?
Query 4: What is the construction technique of cultural property x?
Query 5: What are assets made of 'material x'? What type are they?
Query 6: What are the uses of movable cultural property x?
Query 7: Which entities represent the preparatory phases of cultural good x?
Query 8: What archaeological reconnaissance has been carried out on cultural property x?
Query 9: What assets are part of collection x?
Query 10: What cultural events have affected cultural property x?

4.8 Validation with competence questions

To validate the results of the ontological data consultation, we checked whether the proposed system (SeM) could answer the previously established competence questions.

This phase involved a sample of 28 industry experts who, unaware of the purpose of the experiment, interacted with the developed system by expressing their degree of satisfaction according to the system's answers.

The queries formulated in SPARQL Query Language Schmidt et al. (2010) are shown in Table 3. For each question, domain experts expressed their degree of satisfaction using a questionnaire based on the Likert scale, to which five possible responses were associated: Totally Disagree–TD, Disagree–D, Undecided–U, Agree–A, Totally Agree–TA. The responses obtained are shown in Fig. 5.

As shown in Fig. 5, the degree of user satisfaction is high on average. In detail, users Agree or Totally Agree with a percentage consistently above 75% and, on average, about 80%. Queries 3 and 8 reached about 90% approval due to their specificity and usefulness in the response, which appreciatively impressed users. The degree of Undecided responses reached a considerable percentage due to the diversity of expertise of the selected sample of users.

5 Conclusion

This work aimed to improve the technique for retrieving information from our database for preserving data on the cultural heritage of Campania (DatabencArt).

The aim was to make the data in the platform usable in a way that would be useful to expert users in the field (archaeologists, art historians, geologists, etc.).

In our work, we examined the relationships between the schema of our NoSQL database DatabencArt (in ICCD card format) and the formal ontology ArCo. The data in ArCo are structured based on ontological models that reflect the structural analyticity of the ICCD ministerial cards used to describe the cultural property. Precisely because of the consistency with the structural analyticity of the ICCD cards, we chose ArCo ontology to perform a mapping that would allow us to query our data through a semantic layer (SeM).

To validate the results of the ontological data query, a group of experts in the field assessed whether the SeM system could answer the competence questions and express their degree of satisfaction. The resulting feedback from all questions was positive.

Due to its singularity and versatility, the system has been able to support expert users from different fields, which testifies to how the proposed methodology can be applied to other application areas such as geology and history. In

addition, such a system can be applied to all those contexts in which it is necessary to obtain considerable information to make decisions, such as managing smart environments or the healthcare field.

In conclusion, this integration significantly improves the search for information. However, in order to fully express the data contained in DatabencArt, further work is going in the direction of devising and modelling a purposive ontology on data of a specific domain (Casillo et al. 2022) such as those of the Urban Archaeological Park of Naples; allowing an even more effective and timely consultation of the valuable cultural heritage on the platform.

Funding Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bates MJ (1989) The design for browsing and berrypicking techniques for the online search interface. *Online Review* 13(5):407–431
- Bates MJ (2005) *Berrypicking*. American Society for Information Science and Technology, Medford, NJ, USA, pp 58–62
- Capuano N et al (2015) A personality based adaptive approach for information systems. *Comp Human Behav* 44:156–165
- Carriero, Valentina Anita, et al. 2019 "ArCo: The Italian cultural heritage knowledge graph. In: International Semantic Web Conference. Springer, Cham
- Casillo M et al (2022) An ontology-based chatbot to enhance experiential learning in a cultural heritage scenario. *Front Artif Intellig*. <https://doi.org/10.3389/frai.2022.808281>
- Cole C (2000) Inducing expertise in history doctoral students via enabling information retrieval design. *Library Quarterly* 70(1):444–455
- Corneville F, De Santo M, Dragon M, Gallo L, Troiano A (2020) "DatabencArt and EDUBBA: digital infrastructures for cataloguing and sharing cultural heritage content," in IOP conference series: materials science and engineering. *Int Conf Florence Heri-Tech*. <https://doi.org/10.1088/1757-899X/949/1/012073>
- Cullot N, Ghawi R, Yétongnon K (2007) Db2owl: a tool for automatic database-to-ontology mapping. *SEBD*. 7:491–494
- Colace F et al. (2015) How to Describe Cultural Heritage Resources in the Web 2.0 Era?. In: 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2015 p. 809–815
- Giallonardo, Ester, et al. An architecture for context-aware reactive systems based on run-time semantic models. No. e27702v1. *PeerJ Preprints*, 2019.

- Gruber, Tom. "What is an Ontology? 1993." URL <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>.
- Hau J, Lee W, Darlington J (2005) A semantic similarity measure for semantic web services. In Workshop of WWW2005, Web Service Semantics: Towards Dynamic Business Integration
- Hazber, Mohamed AG, et al. (2015) An approach for mapping relational database into ontology. In: 2015 12th Web Information System and Application Conference (WISA). IEEE <https://paun.databenc.it>. Accessed Apr 2023
- <http://wit.istc.cnr.it/arco>. Accessed Apr 2023
- ICCD, <http://www.iccd.beniculturali.it/index.php?it/1/home>. Accessed Apr 2023
- Jacco van Ossenbruggen, Alia Amin, Lynda Hardman, Michiel Hildebrand, Mark van Assem, Borys Omelayenko, Guus Schreiber, Anna Tordai, Victor de Boer, Bob Wielinga, Jan Wielemaker, Marco de Niet, Jos Taekema, Marie-France van Orsouw, and Annemiek Teesing. Searching and Annotating Virtual Heritage Collections with Semantic-Web Techniques. In: J. Trant and D. Bearman, editors, Proceedings of Museums and the Web 2007, San Francisco and California, March 2007. Archives & Museum Informatics
- Kalfoglou Y, Schorlemmer M (2003) Ontology mapping: the state of the art. *Knowl Eng Rev* 18(1):1–31
- Krabina B (2023) Building a knowledge graph for the history of vienna with semantic mediaWiki. *J Web Semant* 76:100771
- Lora Aroyo, Natalia Stash, Yiwen Wang, Peter Gorgels, and Lloyd Rutledge. 2007 Chip demonstrator: Semantics-driven recommendations and museum tour generation. In: Proceedings of the 6th International Semantic Web Conference (ISWC 2007), LNCS 4825, pages 879–886. Springer.
- Maiga C et al (2020) Building ontology-driven tutoring models for intelligent tutoring systems using data mining. *IEEE Access* 8:48151–48162
- Marjolein van Gendt, Antoine Isaac, Lourens van der Meij, and Stefan Schlobach. 2006 Semantic Web Techniques for Multiple Views on Heterogeneous Collections: A Case Study. In: Proceedings of the 10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2006), LNCS 4172, pages 426–437. Springer, USA
- Moraitou E et al (2019) Semantic bridging of cultural heritage disciplines and tasks. *Heritage* 2(1):611–630
- Moraitou, Efthymia, Yannis Christodoulou, and George Caridakis. "Semantic models and services for conservation and restoration of cultural heritage: A comprehensive survey." *Semantic Web Preprint*: 1–31.
- Ranjgar B et al (2022) An ontological data model for points of interest (POI) in a cultural heritage site. *Heritage Sci* 10(1):1–22
- Rasmussen EM (1999) Libraries and bibliographical systems. ACM Press, New York, NY, Modern Information Retrieval USA
- Schmidt M, Meier M, Lausen G. Foundations of SPARQL query optimization. In: Proceedings of the 13th international conference on database theory; 2010. p. 4–33.
- Shneiderman B (1996) The eyes have it: A task by data type taxonomy for information visualizations. In: Priya D (ed) *IEEE Visual Languages*. College Park, Maryland, USA, pp 336–343
- Sir M, Bradac Z, Fiedler P (2015) Ontology versus database. *IFAC-PapersOnLine* 48(4):220–225
- Uschold M, Gruninger M (2004) Ontologies and semantics for seamless connectivity. *ACM SIGMOD Rec* 33(4):58–64

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.