A unified approach for fast Cultural Goods on site documentation and support multidisciplinary stakeholders fighting illicit trafficking

Charalampos Georgiadis[1[0000-0001-5742-8049]] , Petros Patias[1[0000-0002-2309-085X]], Themistocles Roustanis[2], Haris Zacharatos[3[0000-0002-0164-5607]], Neofitos Vlotomas[3]

[1] The Aristotle University of Thessaloniki, Thessaloniki, Greece
[2] KIKLO IKE, Thessaloniki, Greece
[3] CELLOCK LTD, Nicosia, Cyprus
harrisg@auth.gr

Abstract. This study presents a unified approach to the rapid on-site documentation of cultural goods, tailored to support multidisciplinary stakeholders, including Law Enforcement Agency (LEA) officers and archaeologists, in the fight against illicit trafficking that was developed during the implementation of the EU funded project ENIGMA (DOI:10.3030/101094237) under GA 101094237. The framework emphasizes the use of standardized data models and vocabularies to ensure consistency, interoperability, and efficient data exchange across diverse systems and stakeholders. At its core, the proposed methodology enables LEA officers to conduct fast, structured documentation of cultural goods directly at the site. Leveraging advanced tools, officers can capture critical data, including descriptive information, photos and metadata, which is seamlessly integrated into a centralized system. This data is immediately accessible to expert archaeologists and other specialists, allowing for near-real-time analysis, provenance verification, and enhanced decision-making.

By combining rapid on-site documentation with standardized data structures and real-time collaboration, this approach addresses key challenges in cultural goods management, including data fragmentation and delayed processing. The system not only accelerates the identification and protection of cultural heritage but also fosters effective cross-disciplinary collaboration, representing a significant advancement in efforts to combat the illicit trafficking of cultural goods. All these are provided under a unified digital design and implementation to assist archaeologists and Law Enforcement Agency (LEA) personnel in the methodical examination and recording of cultural goods (CGs). This application streamlines workflows and decision-making processes by combining several capabilities into a single web-based user interface. In this context the LEA personnel initiate the operational framework by performing initial screenings, which include item scanning, document verification, and database searches for any connections to illegal trafficking. Cases are given to experts, including archaeologists, for further investigation, and they add detailed studies to the original data.

Keywords: Rapid on site documentation, Structured data, Cultural goods, Standardized data models.

1 Introduction

Illicit trafficking of cultural heritage objects is a major challenge for law enforcement agencies (LEA) and the archaeological community. The need of the different stakeholders involved in identifying unknown cultural heritage objects and determining their provenance, combined with the need for their rapid documentation, requires tools and techniques to speed up systematic documentation and verification procedures. At the same time, the lack of uniform data structures and the existence of many sources of documentation of cultural heritage objects with different characteristics and terminology between them does not allow for the search for similar objects to guide the researcher on the provenance of an unknown identified cultural heritage object. The EU-funded ENIGMA project (GA 101094237) focuses on endorsing safeguarding, protection, and provenance management of cultural heritage objects. In this context, involved parties such as law enforcement agencies officers and archaeologists need to study and document the cultural heritage objects to identify their provenance.

This paper presents an integrated approach to the documentation of cultural objects when they are discovered in the field by law enforcement agencies and when they are collaborating with their expert counterparts, archaeologists or museum curators, using digital tools. The developed framework emphasizes the critical role of standardized data models and vocabularies in maintaining consistency in the way objects are documented, ensuring interoperability and efficient data exchange between different agencies. In this context data warehouses as a central infrastructure to support a community of multicultural members are based on digital documentation standards that may vary depending on the specific type of cultural heritage object, such as 2D objects such as paintings and manuscripts, and 3D objects such as sculptures, arrowheads, coins etc. The most popular standards that are widely recognized and used in the documentation of cultural heritage artefacts with different levels of detail and concepts such as Dublin Core, Europeana Data Model (EDM), CIDOC CRM, METS, and PREMIS, were reviewed. The final approach focused on using the EDM as the main schema while enriching some structural parts with CIDOC CRM concepts and incorporating Getty vocabularies. This approach was reinforced by the fact that the collaborative cloud data space builds on Europeana's accomplishments in open data and data aggregation. In this way future development can be supported.

The main database infrastructure is the backbone of a digital ecosystem that provides law enforcement agencies with the ability to initially record information on a potential illicitly trafficked cultural object in an easy and fast way and provide the collected information in near real-time to the relevant experts for further investigation. Following the initial object documentation, the experts can proceed to further research and document the object in more detail. In this workflow, stakeholders have several tools at their disposal that provide information on whether the object exists in stolen object databases, tools that can provide suggestions of similar objects to guide the expert on the possible origin of the object if it does not come from a known area.

In this context the described implemented approach is incorporated into the ENIGMA platform that allows a single user interface to integrate information and functions in a way that facilitates decisions about the provenance of the object and whether

or not, it has been stolen. More specifically, this approach leverages digital tools for documentation, research and knowledge acquisition to support rapid and structured documentation, confirming the provenance of objects by optimizing the way experts collaborate in real time. Efficiency is maximized through the proposed solution, as processes are streamlined using workflows, incorporating AI/ML algorithms that assist archaeologists in identifying the object and making the final decision.

The primary advantage of the proposed approach is that this makes the identification and protection of cultural assets faster and enhances efficient inter-site collaboration using a single digital infrastructure.

2 Databases and Standardization

Databases that are used as the central infrastructure to support a community of multicultural members require standardization. Furthermore, digital documentation of cultural assets may vary depending on the specific type of object, for example 2D objects such as paintings and manuscripts, and 3D objects such as sculptures. The data that is collected to document these has many different meanings and embodiments in cultural heritage. In this context they can be categorized in visual and non-visual data.

Representing the physical nature of cultural heritage assets (works of art, architectural works, tools, etc.) is a specific need in the field of cultural heritage. Digital representations that can encode representations of the shape and appearance of the assets are needed. These types of data are called visual data to emphasize the immediate and intrinsic visual copy of the data content.

Representing cultural heritage objects is about more than simply encoding their physical nature. Other information that can be collected, presented, and linked to the asset is called non-visual data. Data that can help characterize materials is an example.

Many types of visual data are used extensively in Cultural Heritage. Main subclasses of visual data are the following: Standard 2D Images, Multi-spectral images, Reflection Transformation Images (RTI), Panoramic Images, 3D representations, Terrain Models, CT data, Videos, Beyond Visual – Sound, and Data produced by diagnostic devices.

Non-visual data and metadata are widely used to describe various characteristics of cultural heritage objects.

In a typical cultural heritage application on the web, data from legacy databases are transformed into a linked data knowledge graph [1]. It includes not only the (meta)data, but also ontologies that define the data modes used (e.g., Dublin Core or CIDOC CRM) and the concepts used to describe the contents (e.g., places, persons, times, and subject matter). The linked data is published using linked data principles and best practices of the W3C [2] in a SPARQL endpoint that facilitates re-use of data in research and application development [3].

The following list presents some of the main standards that are widely recognized and used in the documentation of visual and non-visual data and metadata of the cultural property in Europe and around the world.

- **Dublin Core Metadata Initiative** (DCMI): Dublin Core is a simple and widely used standard for describing digital resources, including cultural heritage items. It

provides a set of 15 core metadata elements, such as title, creator, date, and format, that can be used to describe a variety of resources, and it is widely used by libraries, archives and museums for interoperability [4].

- **Europeana Data Model** (EDM): EDM is a Europeana-specific data model designed to represent the complex relationships between different types of cultural heritage objects and their metadata. It is used and developed by the Europeana foundation to aggregate and share digital cultural heritage data across Europe. The EDM enables integration of diverse cultural heritage datasets and supports linked open data and semantic web technologies [5].
- **CIDOC-CRM** (Conceptual Reference Model): CIDOC-CRM is an international standard for cultural heritage documentation, maintained by the International Council of Museums (ICOM). It provides a conceptual framework for describing cultural objects, events, and the relationships between them [6].
- **LIDO** (Lightweight Information Descriptor Object): LIDO is a standard for describing cultural and artistic objects. It was developed to provide a common format for sharing metadata about museum and gallery collections [7].
- **METS** (Metadata Encoding and Transmission Standard): METS is a standard for encoding descriptive, administrative, and structural metadata about digital objects. It is often used in the digitization and preservation of cultural heritage documents [8].
- **PREMIS** (Preservation Metadata Implementation Strategy): PREMIS is a standard for metadata used in digital preservation. It is essential to ensure accessibility and long-term preservation of digital cultural heritage objects [9].
- **RDA** (Resource Description and Access): RDA is a set of instructions and guidelines for describing and cataloging cultural materials, including books, manuscripts, and other print and digital resources. It provides detailed rules for creating bibliographic records [10].
- **Getty Vocabulary**: The Getty Vocabulary program provides authoritative, vetted vocabularies that provide information about art, architecture, and cultural heritage. These vocabularies are often used to describe and index cultural materials [11].

The choice of standards may depend on the specific needs and type of cultural heritage object being recorded as well as organizational and national requirements.

To support all the process of recording in standardized format and provide a novel documentation mechanism to the involved stakeholders providing them with the ability to compare the objects under investigation with other objects in different sources, two main databases were designed and created, in the context of supporting the ENIGMA platform for fast documentation and the interconnection with open cultural heritage catalogs.

The system architecture is built around a Relational Database designed to facilitate the documentation of essential details about cultural objects gathered by LEA officers. It includes comprehensive general information as well as in-depth provenance data provided by experts. Additionally, it features the primary ENIGMA graph database, which consolidates information obtained from various other sources. This framework enables

end users to analyze and compare the documented details of a cultural object against others within the graph database, allowing them to identify potentially similar objects.

2.1 The Europeana Data Model as the basis of the database management schema

To support the documentation of the cultural objects a database schema based on the Europeana data model was developed. The (EDM) Schema is a flexible framework designed to facilitate the aggregation, integration, and sharing of cultural heritage metadata across Europeana's digital platform. As a key component of Europeana's mission to provide universal access to Europe's cultural heritage, EDM enables the representation of complex relationships between various types of cultural heritage objects and their associated metadata. This model facilitates the interoperability of diverse data sources, enabling museums, archives, and libraries to contribute their digital collections in a structured and consistent manner.

Since the objects documentation involves LEA officers and experts that are not completely familiar with the ontology-oriented schema of Europeana the design of the database aimed on a simplified approach that focuses on the seven most important classes for cultural object documentation. The three core classes of the EDM (**Fig. 1**) are the following:

1. the "provided cultural heritage object" itself (a painting, a movie, a music score, a book…) (edm:ProvidedCHO)
2. one or more accessible digital representations of this object, some of which will be used as previews (the digital picture of the painting.) (edm:WebResource)
3. an aggregation to represent the result of this provider's activity. (ore:Aggregation)

The first two allow capturing the distinction between "works", which are expected to be the focus of users' interest, and their digital representations, which are the elements manipulated in information systems like Europeana.

The third, following the ORE approach, demonstrates that the provided object, together with the digital representations from one Europeana data provider can be regarded as one logical whole.

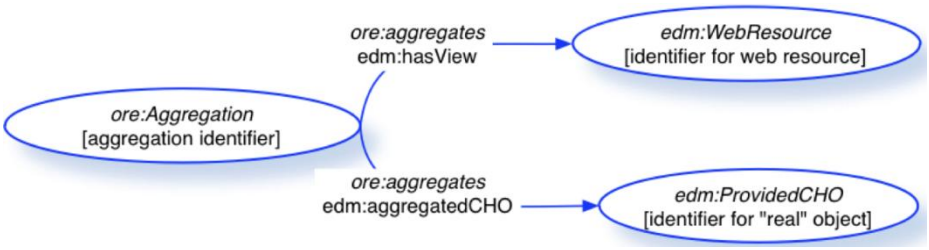

Fig. 1. Europeana's 3 core classes (Source: Europeana Data Model Primer, ttps://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Primer_updated.pdf)

The core five contextual classes are the following: edm:Agent – who; edm:Place – where; edm:TimeSpan - when; skos:Concept – what; cc:License – access and usage.

By leveraging entities such as ProvidedCHO (Cultural Heritage Object), WebResource, Agent, Timespan, Place, and Concept, EDM ensures comprehensive and nuanced descriptions, enhancing the discoverability and contextual understanding of cultural heritage items for researchers, educators, and the general public.

The above-mentioned entities are the main components that were the basis to be used in both relational (database used from stakeholders to record unknown trafficked objects) and main graph database (database that collects information from several sources). The overall list and a short description of how these classes where used is the following:

1. **ProvidedCHO** (Provided Cultural Heritage Object): This entity represents the actual cultural heritage object (CHO). This can be any type of cultural heritage item, such as a coin, statue, manuscript, artefact etc. The ProvidedCHO contains essential descriptive metadata about the object.
2. **WebResource**: Represents the digital representation of the ProvidedCHO. This can be a digital image, video, audio recording, or any other digital file that depicts or represents the cultural heritage object. WebResource entities include information about the format, size, and technical characteristics of the digital files.
3. **Agent**: An entity representing individuals or organizations related to the cultural heritage object. Agents may include creators, contributors, publishers, and other entities involved in the creation, curation, or dissemination of the CHO.
4. **Timespan**: This entity is used to represent temporal information related to the CHO. It may include dates and periods that are significant to the object's creation, discovery, or any other relevant event.
5. **Place**: This entity represents geographic information related to the CHO. This entity can include locations significant to the object's provenance, discovery, or contextual relevance.
6. **Concep**t: An entity representing abstract concepts related to the CHO. These can include themes, subjects, genres, or any other conceptual classifications that help in understanding the cultural and contextual significance of the object.
7. **Event**: Represents actions or occurrences related to the CHO, such as creation, modification, or exhibition. Events provide a detailed context for how the object has been engaged over time.

The present version of EDM integrates the Dublin Core properties from Europeana Semantic Elements, by re-contextualizing each element in the more structured context of EDM. The rationale for the integration is twofold, it maintains unambiguous backward compatibility with ESE and enriches EDM with a set of properties that have proven to be most useful in modelling cultural heritage objects [12].

The values in these properties relate to the original cultural heritage object (CHO) itself, not the digital representation of it. E.g. the attributes of the Mona Lisa, not the digitized image of the painting.

Some of these properties are mandatory to describe better the ProvidedCHO: dc:title or dc:description; dc:language for text objects; dc:subject or dc:type or dc:coverage or dcterms:spatial or dcterms:temporal.

The contextual classes of edm:Agent, edm:Place, edm:TimeSpan and skos:Concept are provided to allow these entities to be modelled as separate entities from the CHO with their own properties if the data can support such treatment. When source metadata contains additional details about an entity that is distinct from the CHO (for instance, the date of birth for an author, different language versions of a subject term) then these contextual classes can be employed to model that separate entity. Such values are properties of that separate entity, rather than the property of the CHO and the properties can be mapped by creation of an appropriate EDM contextual entity. This is the case where the value in the property is an identifier taken from a thesaurus or authority file which will link to further information related to that entity. For example, the identifier for an author name in an authority file will give access to additional information about that author [13].

All the above-described entities in combination with relationships, properties and metadata support the documentation process of the cultural goods.

The main ENIGMA graph database incorporates the RDF schema of the EDM. In this context the main ENIGMA graph database is managing the data based on the providedCHO and creates separate nodes for concepts, places and other contextual information. In order to visually represent this ontology-oriented structure, the following example (**Fig. 2**) represents the corresponding approach [13].

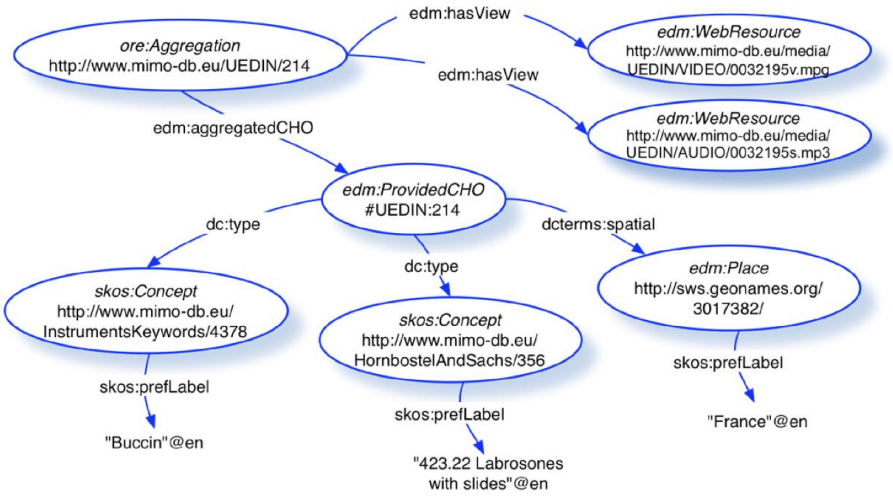

Fig. 2. Representation of the providedCHO and an example of the included contextual classes (Source: https://europeana.atlassian.net/wiki/spaces/EF/pages/987791389/EDM+-+Mapping+guidelines)

To implement the relational database that supports the data collection from LEA officers and experts in the fast documentation process the corresponding relational schema was designed and implemented (**Fig. 3**). The ontology-oriented schema transformed into a relational form creating the necessary relationships between the main entities and incorporating the type of relationship in each one of the relationships. The main entities that were used to emulate the classes of the ontology oriented EDM are the same as described above.

This standardized approach provides the interoperability needed between the two databases and provides the ability to further investigate correlations between objects.

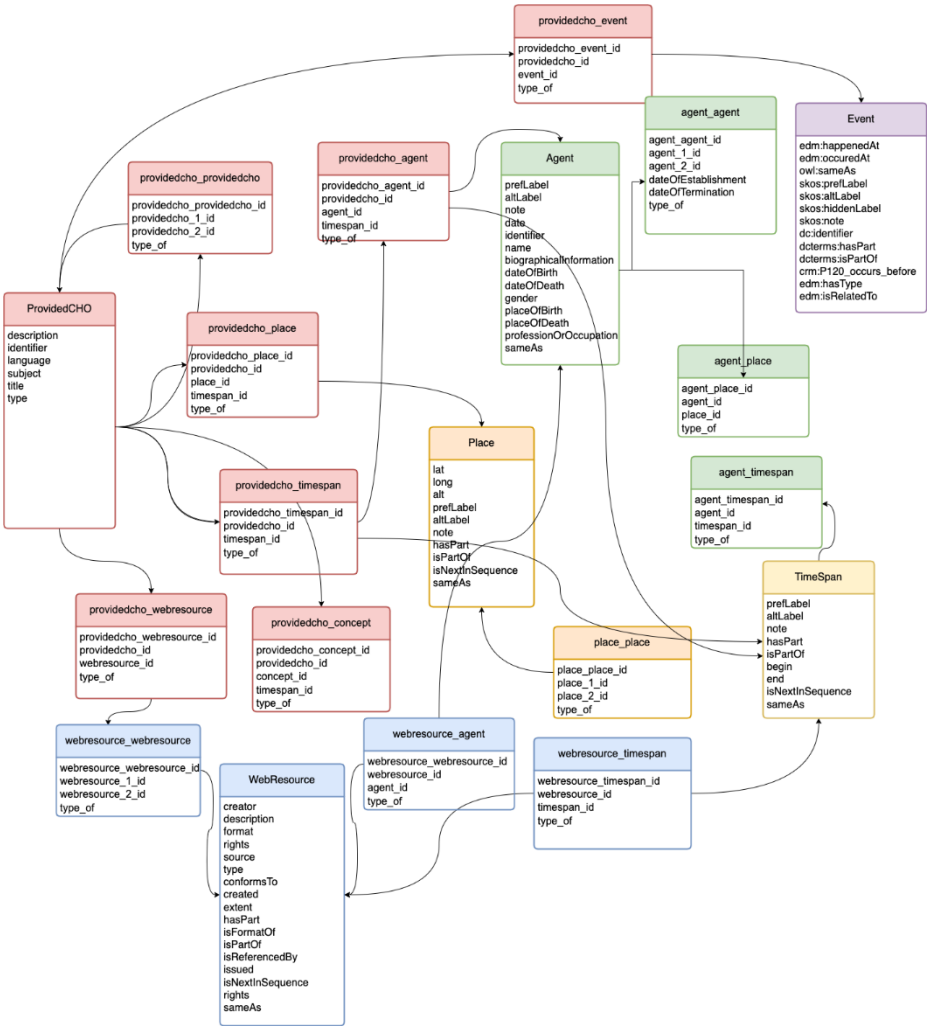

Fig. 3. Abstract representation of the relational database model

2.2 Standardized Vocabularies to support multidisciplinary terminology

The majority of current stakeholders in cultural heritage possess extensive data derived from their engagement with authority files, controlled vocabularies, and thesauri. By integrating contextual resources, it becomes possible to utilize this substantial data while keeping information about the contextual resource distinct from that of the item being described.

For instance, a contributor to Europeana has the ability to create an instance of an edm:Agent class. Instead of merely presenting the text "William Shakespeare" as dc:creator, they can provide a link (URI) to Shakespeare within an authority file, thereby enabling access to rich related information contained in that source, such as multilingual variations of his name and relevant dates and locations of birth and death. This approach applies equally to entities like Places, Timespans, and Concepts.

In developing both the relational database and the primary ENIGMA Graph Database for contextual information, controlled vocabularies were utilized along with URIs to retrieve additional details from external sources when available as linked open data. This method is crucial for enriching the data by incorporating information that may not be present in the initially supplied content.

A variety of controlled vocabulary resources have been examined, leading to the establishment of corresponding infrastructure for their integration into a rapid documentation interface. As an initial measure, the following Getty vocabularies were included providing structured, authoritative terminology for cultural heritage documentation [11]:

- Art & Architecture Thesaurus (AAT): is a thesaurus containing generic terms in several languages, relationships, sources, and scope notes for agents, work types, roles, materials, styles, cultures, and techniques related to art, architecture, and other cultural heritage
- The Union List of Artist Names (ULAN): includes names, rich relationships, notes, sources, and biographical information for artists, architects, firms, studios, repositories, and patrons, both individuals and corporate bodies, named and anonymous. It may include coreferences to other resources to allow interconnections between related disciplines
- The Getty Thesaurus of Geographic Names (TGN): is a unique thesaurus that is global in scope but not comprehensive, focusing on places relevant to art, architecture, and related disciplines. Included are place names, rich relationships, place types, dates, notes, and coordinates for historical and current cities, nations, empires, archaeological sites, lost settlements, and physical features tailored to our core audience. TGN has coreferences to GIS and other resources

These vocabularies enable consistency in data entry and facilitate accurate metadata linking, which is essential for reliable provenance tracking. Once the database schemas have been standardized, it is important to standardize the terminology, as the vocabulary used across museums, archives and libraries is not unified. Furthermore, the application provided to LEA officers and archaeologists for fast documentation incorporates

the ability to search on controlled vocabularies resources in order to have a unified way to describe the corresponding properties of cultural objects.

The vocabulary standardization in this approach is focused on the information that the LEA officers and the corresponding archaeologists/experts provide. An indicative part of the properties that are incorporated into the system to support the vocabularies standardization are described based on the AAT, TGN and the CIDOC CRM categorization:

E57 Material (**AAT Terminology**) (https://cidoc-crm.org/entity/e57-material/version-5.0.2). AAT: The AAT offers a detailed vocabulary for materials, categorizing them into broad classes such as wood, metal, stone, and more specific types like oak, bronze, or marble. In CIDOC CRM E57 Material class encompasses all substances used in the creation of cultural heritage objects. By linking AAT material terms to E57 Material, institutions can standardize the description of an object's composition.

E29 Design or Procedure- Techniques (**AAT Terminology**) (https://www.cidoc-crm.org/Entity/e29-design-or-procedure/version-6.2.1). AAT: Techniques in the AAT are described in terms like carving, casting, painting, etc., providing a structured vocabulary for the processes and methods used in creating cultural heritage objects. In CIDOC CRM E29 Design or Procedure class includes techniques and methods used in the creation of objects. By associating AAT technique terms with E29, the standardization of artistic and manufacturing processes is achieved.

E4 Period (**AAT Terminology**) (https://www.cidoc-crm.org/Entity/e52-time-span/version-6.2.1, https://cidoc-crm.org/Entity/e4-period/version-6.1). AAT: Dates and periods are described using terms such as Renaissance, Baroque, or specific years (e.g., 1500 AD). These terms help categorize objects temporally. In CIDOC CRM E52 Time-Span class represents temporal information, including dates and periods. Associating Getty terms with E52 ensures accurate temporal documentation. E4 Period class is used for historical periods. Linking Getty period terms to E4 helps standardize the categorization of objects by their historical context.

E53 Place (**TGN terminology**) (https://cidoc-crm.org/Entity/e53-place/version-6.2). TGN: Locations are described using hierarchical terms from broad regions to specific sites (e.g., Europe, Italy, Rome). In CIDOC CRM E53 Place class includes all geographic locations. By linking AAT location terms to E53 Place, institutions can standardize geographic information, enhancing the spatial documentation of cultural heritage objects.

3 Fast documentation application

The database infrastructures whose standardization is described in this publication constitute the central support infrastructure of the ENIGMA platform, supporting end users in the documentation of cultural objects and in the identification of their provenance. The ENIGMA platform integrates in a single application the different technologies used in the described framework. The following is a description of the architecture of the application as well as an illustrative mode of operation with the integration of tools to support the end users and ensure the preservation of the described standards.

3.1 Architecture

The architecture that has been implemented allows the communication of the end-user interfaces to the relational database which as described has been designed using the EDM as a basis and at the same time supports the incorporation of the values selected from standardized vocabularies in the corresponding properties. In parallel it supports the enrichment of information from all those hierarchically provided by the linked open data (LOD).

At the same time, this architecture allows for interfacing the applications with the main ENIGMA graph database which as a central repository of data provides information on objects with common features with those entered by the user for the cultural object under investigation. The data flow and the way this architecture provides interaction between the different system components are visually presented in the following schema (**Fig. 4**).

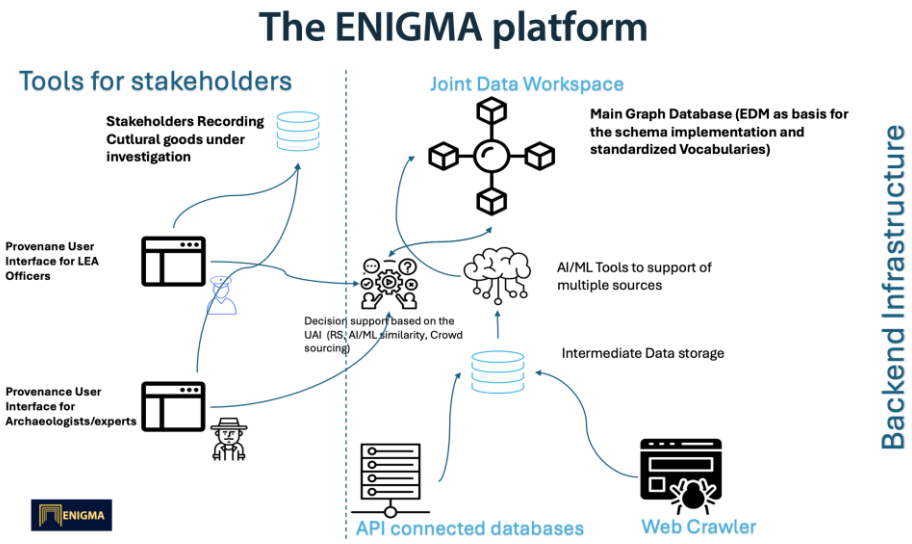

Fig. 4. ENIGMA platform architecture overview

In parallel several other components complete the overall integrated system. The proposed documentation system consists of a web-based user interface designed to support both LEA officers and archaeologists/experts.

Components under this study research

Stakeholders recording Cultural goods database: A repository storing metadata, images, 3D models and documentation properties of cultural object recognized as potential illicit trafficking objects.

Main ENIGMA graph Database: A repository storing metadata, images and 3D models resources and historical records.

Role-Based Access web User interfaces: LEA officers perform initial documentation using a Web user interface customized for their needs, while archaeologists/experts use their own web user interface with more properties to refine, validate data, predict the provenance and decide whether the object falls under cultural heritage protection law or not.

Additional Components

AI/ML Engine: Advanced algorithms analyze object characteristics to support extraction of standardized information from unstructured data and similarity matching for predicting provenance.

Web Crawler: web crawler crawls on specific selected open web resources to collect metadata for similar objects from disparate data structures.

Interoperability Layer: Standardized APIs facilitate integration with existing cultural heritage databases.

3.2 Application workflow

The LEA officers and experts after their authentication are getting access to the ENIGMA platform tool under a unified environment. In the left panel the user has access to the tools that his role is entitled to access. In this case the Expert has access to the provenance research tool, to the earth observation analysis results tool, and the crowd sourcing analysis tool (**Fig. 5**).

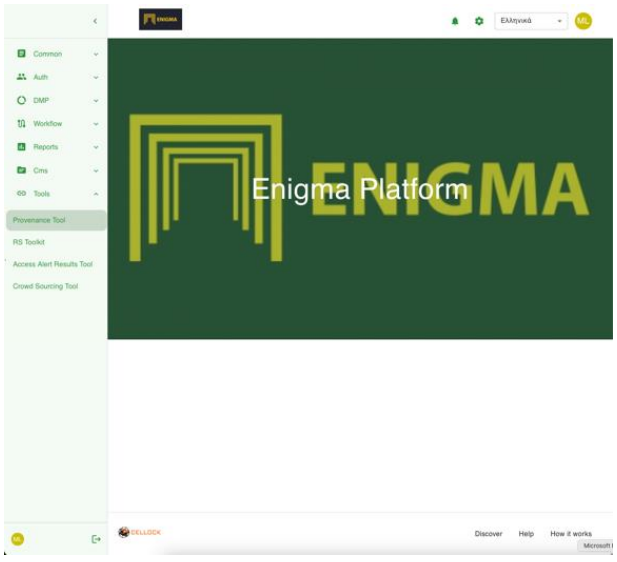

Fig. 5. Navigation panel for Experts access

In the provenance web user interface both the LEA officers and the archaeologists/experts have the ability to review all the assigned cultural objects cases and provide further information to the system using the fast documentation module and get access to the images, 3D content in parallel with additional tools that provide insights based on the similarity investigation on other objects from the main ENIGMA database, AI/ML module and the joint data workspace that provides access to several different sources. Furthermore, access to the earth observation toolkit and crowd sourcing analysis tool provide more information in the field for potential origin areas.

The initial screening of the cultural object is done by the LEA officer by documenting some initial data to define the category, the material, dimensions as well as acquiring images of the object. In this process the tool searches based on this initial and generic information in combination with the images captured for similar objects in the existing registered looted/stolen objects data sources. The interface is enhanced with the ability to search and propose for standardized terminology regarding the properties that have standardized vocabularies. The system proposes the similar values to the LEA officers' input values and also as soon as they select the appropriate value the corresponding hierarchy if exists from the LOD is incorporated (**Fig. 6**).

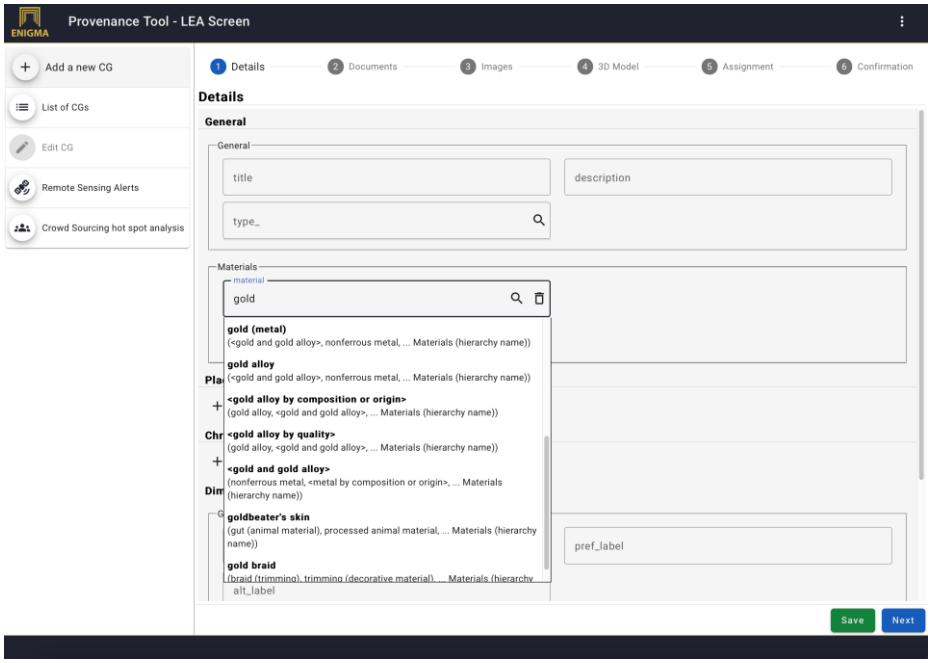

Fig. 6. Indicative example of the documentation user interface for LEA officers and the corresponding Getty vocabularies as LOD integration

The initial screening of the LEA officer is followed by the expert's review that consist of an in-depth analysis and contextual assessment, enrichment of object metadata with historical and geographical insights and cross-checking against several "similar"

cultural objects records from the joint workspace of the system. In this context archaeologists have, through the user interface, additional properties such chronological periods, places of provenance and other more detailed characteristics available to further document and define in a more detailed manner the cultural object identity. All the properties that support standardized vocabularies were enhanced with the same ability (**Fig. 7**).

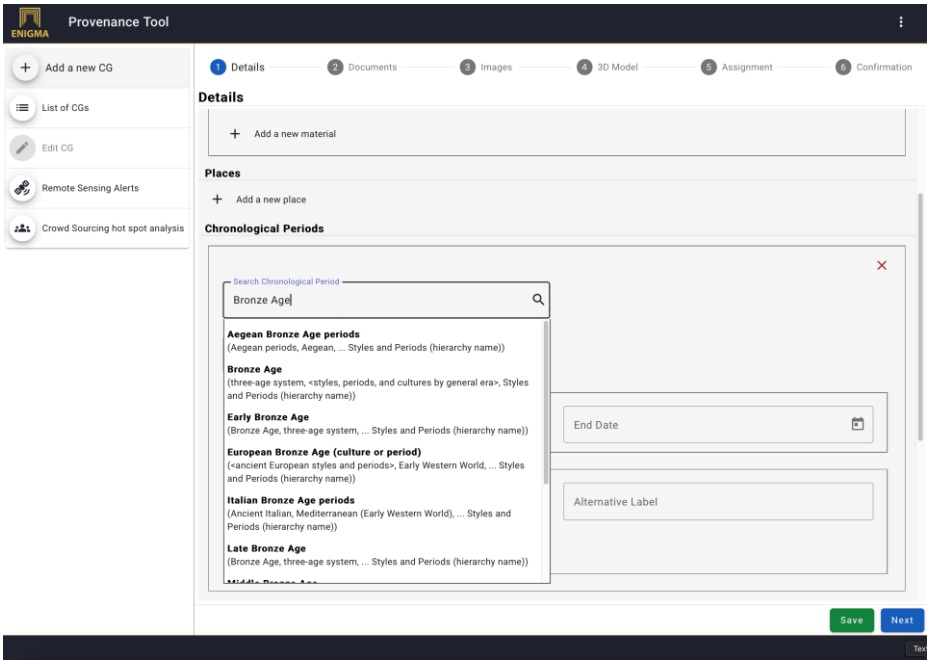

Fig. 7. Indicative example of the documentation user interface for archaeologists and the corresponding Getty vocabularies as LOD integration

4 Conclusions and future work

The implementation of a cohesive digital documentation system greatly improves the effectiveness of cultural heritage preservation. By utilizing standardized data formats, this strategy reduces data fragmentation and facilitates the recognition of provenance. Standardization ensures:

- Interoperability: Consistent data formats enable various stakeholders to easily share and retrieve information.
- Accuracy and Consistency: Defined vocabularies help reduce inconsistencies in documentation.
- Efficiency: Automation of data integration limits manual entry mistakes and speeds up analytical processes.

- Scalability: The system can accommodate extra datasets and technologies as required.

EDM that is used as basis for the design and implementation of both the relational and main ENIGMA graph database allows the enhancement of data from various selected authoritative sources. The origins of such supplementary data are also indicated. This process may:

- Establish connections across languages, domains, and institutional perspectives by aligning with both recognized vocabularies (including individuals, object categories, locations, and time periods) as well as other comprehensive resources like. This creates new meaningful links between diverse objects from different institutions while simultaneously translating metadata through these associations.
- Aid in refining existing datasets by potentially uncovering similar cultural objects records.

The controlled vocabulary integrated into the ENIGMA platform is designed to fulfill the comprehensive requirements of the EDM schema concerning predefined values. The implementation primarily emphasizes Getty vocabularies, which serve as standardized lookup resources for the following reasons:

- Structured, Linked Vocabularies: Includes AAT, TGN, ULAN, CONA, IA.
- Relationships: Incorporates thesauri with three types of relationships: equivalence (synonyms), hierarchical (broader-narrower), and associative (between records that are not hierarchical).
- Standards Compliance: Adheres to established standards for controlled vocabularies (ISO, NISO).
- Comprehensive Knowledge Bases: Goes beyond simple 'value vocabularies'.
- Authoritativeness: Cites credible sources and contributors without being authoritarian; allows for variant terms.
- Scope: Aims to be increasingly multilingual, multicultural, and inclusive within the realm of visual arts.

Under this framework the capability for real-time collaboration between law enforcement agencies (LEAs) and archaeologists represents a significant step forward in combating illegal trafficking.

Future developments should focus on:

- Expanding Database Interconnectivity: Strengthening links with international cultural heritage repositories.
- Developing Mobile Applications: Enabling rapid field documentation via smartphones.
- Implementing Blockchain for Provenance Tracking: Ensuring secure and tamper-proof object identity tracking.

Acknowledgments

This research was funded by the European Union, under the project ENIGMA “Endorsing, safeguarding, protection, and provenance management of cultural heritage” grant number 101094237. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

References

1. Gutierrez, C., Sequeda, J.F. Knowledge graphs. Communications of the ACM 64 (3), 96–104 https://doi.org/10.1145/3418294, (2021).
2. W3C Best Practices for Publishing Linked Data, https://www.w3.org/TR/ld-bp/, last accessed 2025/02/08.
3. Pietroni E. and Ferdani D. "Virtual Restoration and Virtual Reconstruction in Cultural Heritage: Terminology, Methodologies, Visual Representation Techniques and Cognitive Models", Information 2021, Vol. 12, 167: https://doi.org/10.3390/info12040167, (2021).
4. DublinCore, https://www.dublincore.org/resources/metadata-basics/, last accessed 2025/02/08.
5. Europeana Pro, https://pro.europeana.eu/share-your-data/metadata, last accessed 2025/02/08.
6. CIDOC CRM Conceptual Reference Model, https://www.cidoc-crm.org, last accessed 2025/02/08.
7. CIDOC LIDO – Lightweight Information Describing Objects Version 1.1 published, https://cidoc.mini.icom.museum/lido-version-1-1-published/, last accessed 2025/02/08.
8. METS Metatada Encoding & Transmission Standard, https://www.loc.gov/standards/mets/, last accessed 2025/02/08.
9. PREMIS Preservation Metadata maintenance activity, https://www.loc.gov/standards/premis/, last accessed 2025/02/08.
10. Resource Description and Access (RDA), https://www.librarianshipstudies.com/2017/07/resource-description-and-access-rda.html, last accessed 2025/02/08.
11. Getty Vocabularies, https://www.getty.edu/research/tools/vocabularies/, last accessed 2025/02/08.
12. Europeana Data Model Primer, https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf, last accessed 2025/02/08.
13. EDM Mapping guidelines, https://europeana.atlassian.net/wiki/spaces/EF/pages/987791389/EDM+-+Mapping+guidelines, last accessed 2025/02/08.
14. Getty vocabulary program presentation,https://www.getty.edu/research/tools/vocabularies/vocab_what_we_do.pdf, last accessed 2025/02/08.