Article

Bridging "Nature" and "Spirit": The CRMhs Ontology for the Integration of Heritage Science and Cultural Heritage Data

Achille Felicetti * and Francesca Murano

Posted Date: 30 March 2026

doi: 10.20944/preprints202603.2289.v1

Keywords: Heritage Science; ontologies; CIDOC CRM; digital twins; data interoperability; knowledge graphs

Article

Bridging "Nature" and "Spirit": The CRMhs Ontology for the Integration of Heritage Science and Cultural Heritage Data

Achille Felicetti [1,*] and Francesca Murano [2]

1 Fondazione PIN, Piazza dell'Universita, 59100 Prato, Italy
2 Universita degli Studi di Firenze, Italy
* Correspondence: achille.felicetti@pin.unifi.it

Abstract

Heritage Science generates vast quantities of heterogeneous data; however, the absence of a shared semantic framework frequently results in fragmented knowledge and compromised reproducibility. This paper introduces CRMhs, an ontology developed as a formal extension of the CIDOC Conceptual Reference Model (CRM), designed to harmonise the documentation of scientific investigations within the cultural heritage domain. By defining specialised classes for scientific activities, study objects and analytical datasets, the model ensures a robust chain of provenance from initial physical sampling to final interpretative outcomes. The efficacy of CRMhs is demonstrated through its ability to successfully align diverse datasets, ranging from materials-based enquiries to environmental measurements, into a coherent and navigable knowledge graph. This approach not only facilitates seamless data interoperability but also establishes the essential semantic foundation for the advancement of Reactive Heritage Digital Twins. The model bridges the gap between raw scientific evidence and art-historical and archaeological interpretation, fostering a more integrated and sustainable approach to the preservation of cultural heritage.

Keywords: Heritage Science; ontologies; CIDOC CRM; digital twins; data interoperability; knowledge graphs

---

1. Introduction

1.1. Scope and Nature of Heritage Science

Heritage Science emerged as a formally recognised, interdisciplinary domain through a series of landmark institutional interventions. The 2006 UK House of Lords Science and Technology Committee report first catalysed usage of the term, distinguishing Heritage Science from the older 'conservation science' label by stressing its encompassing scope across materials analysis, digital documentation, environmental monitoring, and the social dimensions of heritage [1]. About a decade later, the field had converged on a widely accepted definition of Heritage Science as the interdisciplinary domain of scientific study of cultural and natural heritage, focused on enhancing understanding, care, and sustainable use of heritage so that it enriches people's lives today and in the future. This formulation, now embedded in the funding frameworks of major national and European research programmes, reflects a broad disciplinary consensus rather than any single institutional mandate.

Since then, Heritage Science has been recognised as a distinct, interdisciplinary field that investigates cultural heritage through its material, historical, social, and environmental dimensions. By synthesising methods from conservation practice, the natural sciences, engineering, and the humanities, the discipline provides a robust framework for understanding the complexities of our shared past. The scope of the field is inherently broad, encompassing movable heritage, such as paintings, manuscripts, and archaeological artefacts, as well as immovable structures, including monuments, historic buildings, and archaeological sites. Furthermore, it addresses complex heritage systems such as museum, private collections, archives, and cultural landscapes in which cultural objects are perceived as carriers of aesthetic and historical values as well as physical systems that undergo continuous change over time, interact with their environment and record material evidence of past events and human interventions [2–4].

From a research perspective, Heritage Science addresses a set of fundamental questions that link materials, processes, and meaning. Materials-based enquiries involve the identification of constituent substances, the characterisation of their stratigraphic organisation, and the reconstruction of ancient manufacturing techniques. Process-based questions, conversely, focus on the mechanisms of deterioration and the efficacy of conservation. Researchers seek to identify the physical, chemical or biological drivers of change, assessing how environmental stressors, such as light, humidity, pollutants and biological activity contribute to material risk. This scientific core allows for the development of preventive or interventive measures that are both effective and compatible with the preservation of heritage values. Ultimately, meaning-based questions connect this scientific evidence back to artistic, historical, archaeological and curatorial interpretations. For example, technical imaging might reveal hidden underdrawings or past restorations that fundamentally alter our understanding of an object's attribution and historical narrative [5,6].

1.2. The Heritage Science Workflow

The discipline operates through a sophisticated workflow in which scientific evidence is produced, curated and interpreted across multiple stages. The process typically begins at the documentation stage, where heritage objects are described through cataloguing and condition reporting, frequently enriched with photographic, multispectral and 3D documentation. This is followed by diagnostic investigations that employ a vast spectrum of analytical techniques. These range from non-invasive methods, such as imaging and spectroscopy, to micro-sampling and laboratory-based measurements, all aimed at obtaining detailed material and structural information. These observations are subsequently interpreted within specific conservation and historical frames, leading to informed decisions regarding treatment, preventive strategies and display conditions [5,6].

Crucially, many outputs in Heritage Science are deeply context-dependent. A measured value, such as an elemental composition derived from X-ray fluorescence (XRF), a band assignment in Raman or FTIR spectroscopy, or a humidity time series, only becomes meaningful when it is explicitly associated with its metadata. This metadata must include the object or component to which the measurement refers, the specific activity that produced it (including instrument settings, calibration and sampling strategy), and the interpretative model used by the researcher. For instance, a chemical signature remains a mere datum until it is linked to a specific material identification hypothesis or an ageing mechanism. Consequently, Heritage Science requires formal representations that preserve not only the final results but also their entire provenance, associated uncertainties and the underlying assumptions of the investigative process [6].

1.3. The Challenge of Data Heterogeneity

Heritage Science is inherently data-intensive. However, the data are produced by a wide variety of professional communities, including conservators, museum scientists, chemists, physicists and archaeologists [3]. These groups often use different terminologies, measurement conventions, formats and storage systems, ranging from local collection management systems to laboratory notebooks and sensor platforms. Even when two different projects employ the same analytical technique, they may report their results at different levels of granularity. For example, one study might provide spot analyses while another provides high-resolution maps, each using different descriptions of uncertainty and different conceptualisations of material stratigraphy [2,6].

This heterogeneity becomes particularly critical when research questions require analysis across multiple objects or collections. Such tasks might include comparing pigment usage across different geographic regions and periods, or correlating environmental exposure with observed deterioration patterns on a national scale. Achieving this level of integration requires a semantic layer capable of aligning data not only at a syntactic level but also at a conceptual one. It is necessary to connect observations to objects, parts, samples, instruments and interpretations in a consistent and machine-readable way. This ensures that the data can be queried and compared regardless of its origin or the specific system in which it was initially recorded.

The need for interoperability in Heritage Science extends far beyond the simple harmonisation of field names or file formats. The core problem concerns the ability to relate all aspects of the scientific process to a coherent narrative of what has happened to a heritage object. This implies at least three fundamental requirements for any semantic encoding layer. First, descriptions must be event-aware and activity-aware. Measurements, sampling events and analytical processing are all activities that introduce specific contexts and dependencies. These must be modelled explicitly to enable any meaningful comparison or reuse of the data [2].

Second, the representation must support full provenance and traceability across the entire knowledge lifecycle. This includes the path from the physical object and its components to the individual sample, the raw measurement, and finally to the derived results and interpretative statements. Without this traceability, it is impossible to validate conclusions or to reproduce analytical pipelines.

Third, the approach must accommodate the evolution of the field. Heritage Science projects differ widely in their aims, and the conceptualisations used by researchers evolve as new methods and technologies are developed. A semantic layer should therefore allow for extensions and specific mappings while maintaining a stable core model that supports cross-project integration.

These requirements have motivated the increasing adoption of semantic technologies and community reference models such as the CIDOC CRM. By providing a mechanism to describe entities and processes explicitly, these ontologies allow for the alignment of local vocabularies to shared conceptual structures. This supports rich, complex querying over heterogeneous datasets, transforming raw data into actionable knowledge. This paper introduces CRMhs as a formal response to these requirements, providing a specialised extension designed to capture the unique nuances of Heritage Science investigations. The following sections describe the development of this model and its application to some significant archaeological discoveries.

The paper is organised as follows. Section 2 explores the existing landscape of Heritage Science data management, highlighting the semantic challenges inherent in multidisciplinary research. Section 3 presents the formal architecture of the CRMhs model, detailing its core classes and their alignment with the CIDOC CRM ecosystem to ensure a rigorous chain of provenance. In Section 4, the practical application of the ontology is demonstrated through a series of case studies, illustrating its capacity to harmonise diverse analytical datasets into integrated knowledge graphs. Section 5 discusses the broader implications of the framework, focusing on its capacity to bridge the epistemological divide between raw scientific datasets and high-level cultural interpretations. Finally, Section 6 offers concluding remarks on the modular evolution of the model and outlines future directions for its implementation.

2. Overview of Relevant Models and Methodologies

2.1. The Landscape of Conceptual Models for Scientific Data

The international research landscape offers a diverse array of conceptual models designed to document the complex information generated by scientific investigations. While some frameworks are general in scope, others have been meticulously developed to serve the needs of specific disciplinary communities. However, prior to the development of CRMhs, no conceptual model had been explicitly conceived to address Heritage Science as a distinct and autonomous domain. This is mostly related to the primary challenge of the discipline that lies in the necessity to capture the unique relationship between scientific enquiry and cultural heritage, representing the full complexity of their multi-layered interactions. Heritage Science is uniquely characterised by the integration of rigorous analytical methodologies, typical of the natural sciences, with the interpretative and contextual approaches

 doi:10.20944/preprints202603.2289.v1

rooted in the humanities. Consequently, the design of the CRMhs model was preceded by a systematic examination of existing conceptual frameworks in both the analytical sciences and cultural heritage sectors. This process aimed to draw upon established best practices while combining their respective strengths to ensure a comprehensive representation of the domain.

Within the broader scientific community, several general-purpose models provide a foundation for the semantic representation of research information. For instance, the Common European Research Information Format (CERIF) [7] provides a highly structured metadata model for describing research entities and their interrelations, including an extensive vocabulary of roles and types commonly used in academic and industrial research contexts. Similarly, the Extensible Observation Ontology (OBOE) [8] offers a robust structure for modelling knowledge related to scientific observations. Both models supply essential classes and properties that enable detailed semantic descriptions of scientific activities and their resulting datasets [9,10].

Other initiatives are more domain-specific and have been developed by large scientific communities to systematise discipline-oriented metadata. Notable examples include NeXus [11], a common data format for neutron, X-ray and muon science, and the Astronomy Visualization Metadata framework (AVM) [12], which focuses on astronomical imagery. The Crystallographic Information Framework (CIF) [13] is particularly noteworthy for its modular structure and the rich set of complementary resources it provides, including comprehensive domain dictionaries that enable a high degree of terminological standardisation. Similar terminological resources exist in fields such as mineralogy (see for instance the IMA Mineral List [14] or the Mineralogy Database [15]), while NASA maintains a widely adopted general-purpose thesaurus that serves as a benchmark for aerospace and related sciences [16]. A particularly relevant example is also the CHAracterisation METhodology Ontology (CHAMEO), a domain ontology designed to model common aspects across different characterization methodologies in materials science [17]. CHAMEO is part of a broader initiative under the European Materials Modeling Council (EMMC) to develop interconnected materials modeling ontologies based on the Elementary Multiperspective Material Ontology (EMMO), and is particularly valuable for Heritage Science because it provides a standardized terminology and structure for modeling characterization techniques, facilitating the integration of scientific data from heritage objects into broader materials science research [18].

Many scholars [19] highlight the specific challenges that Heritage Science faces in implementing FAIR data principles, noting that the interdisciplinary nature of the field, where historical and cultural contexts are as significant as analytical methods and material properties, is rarely accounted for by standard scientific data management templates [19]. Curated resources like the INFRA-ART Spectral Library, ensures that material evidence remains accessible and verifiable for the international research community. As a critical digital support tool registered within the European Open Science Cloud (EOSC) portal, this library provides the high-quality, curated spectral data, encompassing over 900 materials across XRF, ATR-FTIR, and Raman spectroscopy, required to populate the complex scientific workflows modelled in this study [20].

Significant international initiatives, such as the European Research Infrastructure for Heritage Science (E-RIHS), are currently instrumental in steering the global discourse toward the definition of shared architectures and data management frameworks [21]. The transition of their activities from overarching roadmaps to the delivery of operational, machine-actionable semantic models, however, currently remains in a predominantly developmental or consultative phase. Recent national initiatives such as the Italian H2IOSC project [22] have also begun to explore the requirements for a dedicated Heritage Science ontology, but the focus was primarily concentrated on preliminary knowledge elicitation and the conceptual definition of domain boundaries [23].

The cultural heritage domain likewise offers significant conceptual and terminological resources. Among these, the CIDOC Conceptual Reference Model (ISO 21127:2023) [24] stands out as the most widely recognised framework for the semantic representation of heritage information. Its ecosystem of extensions provides a flexible structure capable of accommodating a wide range of sub-domains, from

 doi:10.20944/preprints202603.2289.v1

archaeology and museology to epigraphy and bibliographic documentation. Of particular relevance to the present study is the CRMsci extension, which was specifically designed to represent scientific observation and measurement processes. To date, CRMsci constitutes the most comprehensive attempt to integrate scientific and humanistic information within a coherent conceptual structure [25].

CRMhs builds upon this foundational work, producing records that are closely linked to CRMsci and fully compatible with the CIDOC CRM ecosystem. This alignment is essential for ensuring interoperability with major European research infrastructures and initiatives that have adopted CIDOC CRM as the backbone of their digital architectures. By positioning CRMhs within this established semantic environment, it becomes possible to bridge the gap between analytical data and the broader historical and curatorial context of heritage objects.

The semantic efficacy of the CRMhs framework is further enhanced by its alignment with established international vocabularies, which provide the terminological granularity necessary for cross-domain interoperability. Within the Cultural Heritage sphere, the Getty Art & Architecture Thesaurus (AAT) [26] serves as the primary reference for the standardised classification of materials and objects, while PeriodO [27] offers a robust platform for the formalisation of spatio-temporal boundaries and chronological definitions. Crucially, resources such as PeriodO, the Mineralogy Database and other geographic gazetteers such as Pleiades [28], are instrumental in bridging the gap between material characterisation and geological provenance and in establishing verifiable links between geochemical signatures of artefacts and their specific geological point of origin, required for high-resolution provenance studies.

3. The CRMhs Ontology: A Comprehensive Framework for Heritage Science

3.1. Overview of CRMhs Develepement

The CRMhs ontology has been developed to capture and represent the full range of scientific activities involved in Heritage Science. It results from the integration of several existing ontologies that already provide robust models for general aspects of research, while introducing new entities specifically designed to address the distinctive features of scientific investigation within the cultural heritage domain. The development of CRMhs is the result of a rigorous collaboration with prominent scientific research institutions. The preliminary results of this developmental activity, which focus on the initial conceptual framework and the identification of core ontological requirements, are documented in several seminal papers establishing the essential theoretical foundations for the model, specifically addressing the challenges of domain boundary definition and the integration of multi-modal scientific datasets within a unified semantic ecosystem [29–32]. In particular, the model was forged through the investigation of laboratory datasets from the Italian CNR, INFN, and the Opificio delle Pietre Dure, with a meticulous focus on their diverse encoding formats and the specific archiving protocols utilized [33,34], an empirical grounding ensuring that the representation of scientific workflows and structured research data is both technically robust and deeply rooted in actual laboratory practice. By building upon these documented strategies, the current iteration of CRMhs transitions from the preliminary phase of survey and elicitation toward a fully operational, machine-actionable implementation. Additionally, CRMhs has been significantly informed by practical insights and modelling efforts conducted within major European initiatives, notably the ARIADNE [35], 4CH [36] and ARTEMIS [37] projects. ARTEMIS in particular has adopted CRMhs as the reference ontology for the integration of scientific analytical data within its Heritage Digital Twin ecosystem [38,39]. More broadly, several heritage science research laboratories have independently have identified the CRMhs modelling principles as the appropriate framework for the semantic management and publication of their analytical data [40].

The ontology addresses two primary strategic objectives. First, it provides a conceptual framework for institutions such as museums, universities and laboratories that need to organise, document and preserve the results of scientific investigations, particularly those that have not yet established a structured digital archive or laboratory repository. Second, it offers a semantic tool for the standardi-

 doi:10.20944/preprints202603.2289.v1

sation, integration, sharing and reuse of scientific data across different laboratories and institutions. By enabling alignment with the wider cultural heritage domain, CRMhs promotes interoperability between scientific and humanistic research environments. To achieve these goals, the model adopts the data modelling principles of the CIDOC CRM and aligns with the FAIR principles to ensure that digital scientific data remains findable, accessible, interoperable and reusable [2,14,25]. A comprehensive technical description of the CRMhs ontology, encompassing the full hierarchy of classes, properties and formal axioms, is available through dedicated Zenodo [41] and GitHub [42] repositories, and also include the RDF encodings of the model and practical usage examples.

3.2. *Ontological Foundations*

CRMhs adopts the terminological framework of CIDOC CRM [24] and its extensions, utilising the notions of class, subclass, property, domain, range, inversion and inheritance to define its structure. These principles underpin the semantic coherence of the model, ensuring compatibility with international standards. The ontology is organised around a hierarchy of entities, ranging from high-level concepts to specialised classes tailored to the scientific processes of heritage investigation. Relationships between entities are defined through a set of properties that support the documentation of workflows, provenance, analytical procedures and data production.

A distinctive feature and strength of CRMhs lies in its hybrid structure, which combines domain-specific classes developed to capture the peculiarities of Heritage Science with the direct reuse of general-purpose classes and properties from the CIDOC CRM ecosystem. Entities pertaining to common aspects of cultural heritage documentation, such as time, place, actor and physical object, are modelled entirely through the existing semantic framework of CIDOC CRM, without redefinition or duplication. Consequently, CRMhs reuses classes such as *E39 Actor*, *E53 Place*, *E52 Time-Span* and *E77 Persistent Item*. This layered approach ensures conceptual consistency and seamless integration with other CRM-based ontologies, creating a unified conceptual space where scientific and humanistic knowledge can be expressed together. The ontologies integrated within CRMhs are summarised in Table 1, which also reports their versions, namespaces, brief descriptions, and the abbreviations used to identify their classes and properties both in their official documentation and in this paper.

Table 1. Ontological models integrated within the CRMhs framework.

Model	Version	Prefix	Description	Classes	Properties
CIDOC CRM	7.1.3	crm	Cultural Heritage ontology	E	P
CRMsci	2.0	crmsci	Scientific observation and measurement model	S	O
CRMdig	3.2.1	crmdig	Model for provenance metadata	D	L
CRMinf	1.2	crminf	Support for scientific argumentation and inference	I	J
CRMtex	2.0	crmtex	Model for the study of ancient texts	TX	TXP

3.3. *Classes and Properties of the CRMhs Model*

3.3.1. Scientific Activities and Analytical Procedures

Scientific work in this field revolves around a wide range of coordinated activities, some experimental, some preparatory, and others interpretative or supportive. These actions are performed in specific temporal and spatial contexts, following rigorous standards and methodologies. CRMhs models these processes through a coherent set of entities designed to describe both the analytical procedures and the broader operational framework.

To represent the execution of scientific procedures, CRMhs introduces the top-level class HS1 Scientific Activity, designed to represent any kind of operation related to scientific investigation. This includes the handling of objects, the preparation of samples, the operation of devices, the use of software, the collection and processing of data and the calibration and setup of tools. Each activity is embedded in a procedural framework that captures not only what was done, but how, with what, and under which conditions. Activities can be decomposed into sub-activities using the property *HSP3 has sub-activity*, allowing for the representation of complex workflows in discrete, traceable steps.

The execution of an activity is often guided by formalised protocols. The property *HSP7 used protocol* links an activity to the protocol that governed its execution, while *HSP6 used method* connects it to a more general methodological framework. Technical configurations are documented through *HSP8 used device*, *HSP9 used component* and *HSP10 used software*. Furthermore, the model supports the integration of pre-existing data through *HSP12 used dataset* and the generation of new results via *HSP16 produced dataset*. In keeping with the event-centric approach, each activity is temporally and spatially contextualised using *E52 Time-Span* and *E53 Place*, linked via *P4 has time-span* and *P7 took place at*.

The class *HS3 Analysis* is a specific subclass of *HS1* that captures activities aimed at obtaining data through observation or measurement. This encompasses both destructive and non-destructive techniques conducted on entire objects, portions, or samples. Analyses are specifically linked to their subject of study via *HSP4 analysed*. To account for the conditions of a measurement, *HSP13 used settings* links the analysis to one or more *HS14 Analysis Settings* entities, which may carry a value (*HSP14 has settings values*) or be specified in an external dataset (*HSP15 is specified in*). In many cases, analysis is preceded by a *HS10 Sample Preparation*, another subclass of *HS1*, which includes cleaning, mounting, embedding, or other treatments required to make the sample analysable. This is modelled through the property *HSP5 prepared*, ensuring full traceability of material transformations.

3.3.2. Physical and Digital Scientific Objects

Heritage Science is characterised by the interplay between tangible materials and their digital representations. CRMhs provides a comprehensive framework to describe both dimensions and to model their relationships within the research workflow. The class *HS6 Study Object* represents the primary entities selected as the focus of scientific investigation. Defined by their epistemic role, these entities are actively examined through observation, measurement, or experimentation. By being declared as a superclass of *E18 Physical Thing*, *HS6* situates the scientific act of inquiry within the material branch of the CIDOC CRM. This role-based aggregation, analogous to *E72 Legal Object*, ensures that any physical entity may become a study object without altering its ontological status. All information regarding identity, provenance, typology, production and custody continues to be expressed using standard classes such as *E18 Physical Thing* or *E22 Human-Made Object*.

The class *HS9 Study Object Sample*, a direct subclass of *HS6* and *S13 Sample*, represents discrete physical portions extracted for investigation. Unlike an *HS7 Study Object Portion* that remains an integral part of the parent fabric, an instance of *HS9* is a newly individuated entity resulting from an act of separation. These samples are characterised by a constitutive relationship with the object of provenance, distinguishing them from generic specimens. Through the property *HSP17* was taken from, the model ensures a rigorous material traceability chain, preserving information regarding extraction conditions and the specific areas of origin.

3.3.3. Object Portions, Areas and Localisations

Investigations may focus on specific parts of an object or surface areas of interest. *HS7 Study Object Portion* models physically identifiable parts, such as a structural component or a stratigraphic layer. These are often situated within broader regions represented by *HS8 Study Object Area*, which is also a subclass of *E53 Place*. Portions are connected to their areas via *HSP18 is located within*, while the provenance of a sample from a specific area is recorded using *HSP20 was taken from area*, complementing the general link to the parent object (*HSP19 was taken from*). Furthermore, *HSP17 has area of interest* links a study object directly to one or more areas. Each area can be precisely located on the object's surface using *HSP21 has object surface coordinates*, which refers to *E47 Spatial Coordinates*.

 doi:10.20944/preprints202603.2289.v1

3.3.4. Devices, Datasets and Software

HS11 Scientific Device represents the tools employed in procedures, such as spectrometers or microscopes, which can be further decomposed into *HS12 Device Component*. Digital outputs are represented by HS13 Scientific Dataset (subclass of *D9 Data Object*), which captures raw data files, processed results, or interpretative outputs. These are created using *D14 Software*, linked via *HSP10 used software* and *HSP11 was created using software*. Finally, *HS14 Analysis Settings* represents the parameters and calibration values that influence an analysis, ensuring the interpretative reliability of the scientific results [5,16].

3.3.5. Collaborative Context and Actors

Scientific investigations involve multiple actors and institutions, including research bodies, universities, laboratories and technical staff. The documentation of these activities in CRMhs is designed to interoperate with the CIDOC CRM core ontology, which includes classes for actors and roles. Additional participants, such as museums, private collectors, archaeologists and curators, are often involved as holders or discoverers of the objects. While their roles may not be strictly technical, they are essential for framing the scientific process within a broader cultural and institutional context, and their contributions are documented accordingly to provide a holistic view of the heritage investigation. An overview of the classes and properties of the CRMhs model is presented in Figure 1. The complete documentation of the model is available through Zenodo [41] and GitHub [42].

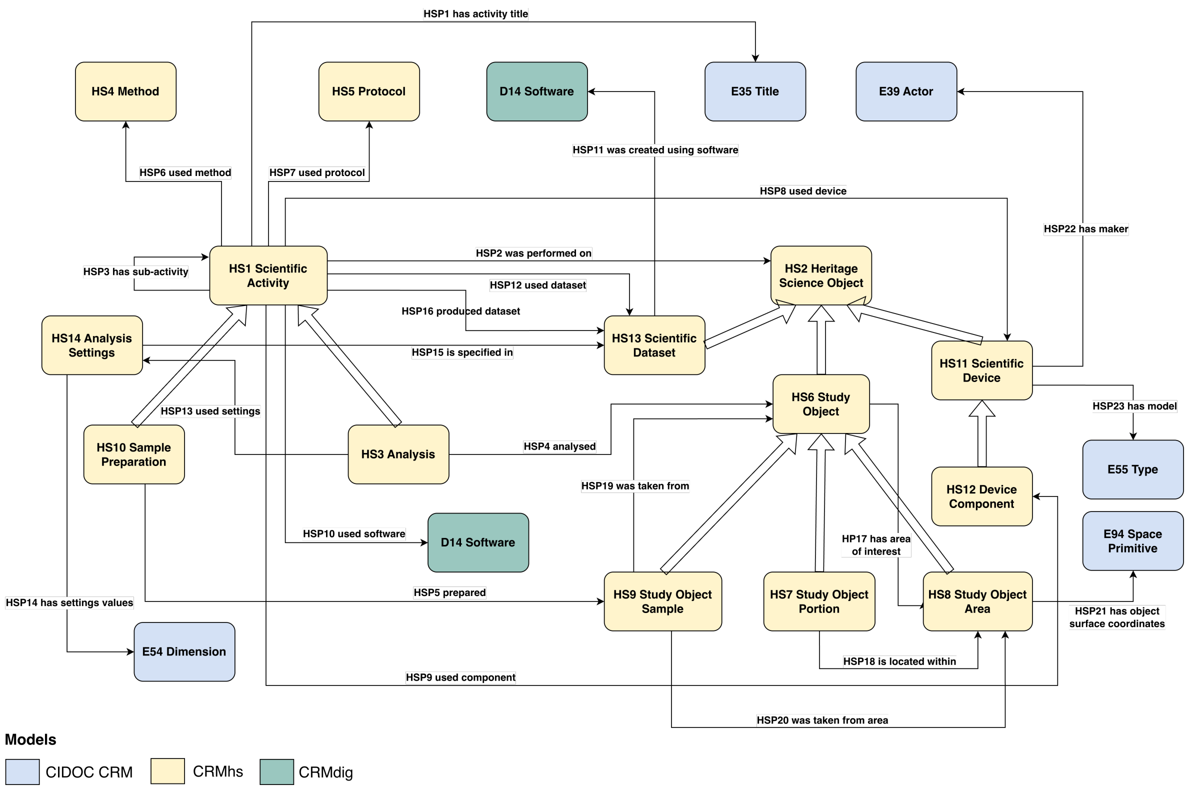

Figure 1. Classes and properties of the CRMhs model.

4. CRMhs in Action

CRMhs has been instrumental in structuring scientific data for cultural heritage in many different contexts. A notable example was accomplished in the ARTEMIS project, where unstructured textual reports documenting Macro X-ray Fluorescence (MA-XRF) analyses of famous paintings were processed using AI pipelines [43]. CRMhs was combined with Large Language Models to identify and extract domain-specific entities (e.g., analysis techniques, materials, instruments) from complex scientific texts, while semantic similarity algorithms resolved synonyms and variants [39,44]. The extracted data were then encoded as RDF triples using CRMhs entities (e.g., *HS3 Analysis*, *HS11 Scientific Device*), to

 doi:10.20944/preprints202603.2289.v1

populate a Knowledge Graph. This AI-assisted workflow transformed raw textual documentation into structured, queryable knowledge, enabling advanced semantic searches (e.g., linking pigments to their spatial distribution in the painting).

Additionally, various CRMhs-based interfaces tailored for laboratory workflows have been developed to simplify scientific data collection, such as the THESPIAN suite designed for INFN-CHNet's heritage science data [33]. THESPIAN offers a metadata form where fields (e.g., study object, analysis device) map directly to CRMhs properties and uses machine learning models to semi-automate entity completion using content from Italian-language reports. These tools, deployed in the 4CH Cloud Platform, ensured FAIR compliance and reduced manual encoding burdens, as demonstrated by their use in documenting INFN-LABEC's XRF analyzes and radiocarbon dating datasets.

The two application examples presented in this paper are situated within the domains of archaeology and epigraphy and are intended to illustrate the potential and utility of CRMhs to model scientific investigations of ancient artifacts. Both case studies focus on epigraphic materials that have undergone scientific analyses on supports aimed at addressing distinct research questions. The results obtained from these analyses have enabled a more comprehensive understanding of the investigated objects, their nature and significance.

The application of the CRMhs ontology to the these case studies allows the entire analytical procedure to be represented as a semantically interconnected knowledge graph. In this framework, each stage of the scientific workflow, from the physical documentation of the artifacts to the generation and interpretation of isotopic and elemental datasets, is formally represented through the integration of CRMhs and the CIDOC CRM ecosystem, a modeling approach that ensures that the relationships between material entities, scientific activities, instruments, datasets and interpretative conclusions remain fully transparent and traceable.

4.1. Magic and Science: Scientific Investigations of Archaeological Artefacts Related to Aggressive Magical Practices

The first case study pertains to materials related to the practice of aggressive magic known as *defixiones* [45,46]. *Defixiones*, or curse tablets, constitute a well-attested textual genre in the ancient world. These inscriptions, typically inscribed on thin lead sheets, contain texts of varying complexity that were "intended to influence, by supernatural means, the actions or welfare of persons or animals against their will" [47]. The corpus of *defixiones* extends across a broad geographical range, from the Mediterranean to Britannia, and covers an extensive chronological period, spanning from the 6th century BCE to late antiquity. This tradition encompasses multiple linguistic contexts, including Greek, Latin, Oscan, Gaulish, Etruscan and Punic, as well as others. The creation of a magical artefact such as a *defixio* involved a complex ritual process. The lead sheets were often rolled, folded, or bound with thread, or pierced by nails, which were occasionally left embedded [48]. As part of the ritual, these tablets were concealed in tombs or in locations associated with the underworld, such as wells or springs [48], or in sanctuaries dedicated to chthonic deities, such as the shrines of Malophoros at Selinus and of Demeter at Corinth.

To further enhance their efficacy, the tablets were sometimes accompanied by objects belonging to the intended victim (*ousía*), such as locks of hair, nail clippings, or fragments of clothing, as well as by figurative representations of the victim. This practice relied on the principle of *similia similibus*, a mechanism also observed in other forms of aggressive magic, such as those found in Vodou rituals. A relevant example of aggressive magic is a lead statuette dating to the second half of the 4th century BCE, which is currently housed in the Antikensammlung der Staatlichen Museen zu Berlin (inv. 30741, https://id.smb.museum/object/680844; [49], see Figure 2). Approximately 90 such objects have been documented from antiquity, primarily made of lead, though other materials, such as wax or clay, were also employed. Among the most renowned examples is the clay figurine from Egypt, pierced by thirteen pins, which is now housed in the Louvre, Paris (Louvre inv. E 13623, https://collections.louvre.fr/en/ark:/53355/cl010044394). Measuring 9.6 cm in height, the lead statuette was cast and subsequently refined with a blade to carve its details. It bears a Greek

 doi:10.20944/preprints202603.2289.v1

inscription listing the names of multiple victims, which suggests that its purpose was to symbolise the act of binding in a general sense rather than to target a specific individual [50].

Figure 2. Statuette 30741 from the Antikensammlung der Staatlichen Museen zu Berlin. *Credit: Staatliche Museen zu Berlin, Antikensammlung / Norbert Franken CC BY-SA 4.0*

This statuette was acquired in 1904 in Greece by H. Schrader, who later sold it to the Berlin museum in 1918. Regrettably, no reliable information regarding its origin has been documented. The only account stems from the antiquarian who sold it to Schrader, who alleged that the statuette originated from Corinth. To investigate the statuette's provenance, an analysis of its lead isotopic composition was conducted using mass spectrometry. By comparing the isotopic composition with that of known lead ore deposits, researchers have been able to determine the geographical origin of the lead used in the statuette's production; this also allowed the reconstruction of the trade routes of processed lead [51].

Lead isotope analysis was conducted on extremely small samples (1 mg of a lead alloy). A small subsample (~1 mg) was initially cleaned with ultrapure water in an ultrasonic bath and then dissolved in ultrapure nitric acid. The dissolved sample then underwent liquid ion chromatography to isolate the lead from any impurities and alloying components. Following purification, two sample filaments were prepared from the lead fraction, and their isotopic composition was determined using thermal ionisation mass spectrometry. To account for mass fractionation, the isotope reference material NIST SRM 981was employed, processed under identical conditions as the sample to validate the analytical procedure. To determine the provenance of the lead in the statuette, a geographical assessment of the lead isotopes was made by comparing it with isotopic compositions in ore deposits from Europe (especially from the Mediterranean area) specifically collected and evaluated. The isotopic composition of the statuette lies within the isotopic variation observed in Aegean lead ores. Given the isotopic overlaps, it is highly probable that the lead used for the statuette originated from either Keos or Laurion, located 60 km SE of Athens, between Thorikos and Cape Sounion [51]. According to Curbera [50], the high concentration of lead curse tablets in Athens is linked to the ready availability of inexpensive lead, which was a by-product of large-scale silver smelting operations at the Laurion mines.

From an ontological point of view, the lead statuette, its traditional provenance statements and the scientific provenance investigation performed by Jochen Vogl and Martin Rosner (*E39 Actor*), are characterized by the integration of physical evidence, historical documentation and scientific reasoning through the formal structures of CRMhs. The figurine itself can be instantiated as an *E22 Human-Made Object*, to reflect its status as a manufactured ancient artifact. This automatically turns it into a *HS6 Study Object*, being *E22* declared as subclass of *HS6* in CRMhs. Its primary material, lead, is documented through the property *P45 consists of* pointing to the class *E57 Material*, while the *P108 was produced by* property can be used to document its production process (*E12*) and place (*E53*).

The original account provided by the antiquarian seller (*E39 Actor*), who alleged the object originated from Corinth, is modeled using a set of entities provided by CRMinf, an ontology specifically designed to document the provenance of knowledge by representing observations, argumentation, inferences and beliefs made by actors regarding specific aspects of reality [52]. In this perspective, the seller's statement (*I4 Proposition Set*) can be modelled as the subject of an *I2 Belief* (and the property *J4 that*) and an *I1 Argumentation* (using the *J2 concluded that* property) on which his claim was based. Interestingly, the proposition set (*I4*) is not just a text fragment, but can be semantically detailed through a series of ontological statements using the classes and properties of the various ontologies. In this case, the *E12 Production* and *E53 Place* classes of CIDOC CRM, linked through the *P7 took place at* property, are used to build the structure of the statement regarding the supposed provenance of the statuette.

Concerning the scientific analysis, the extraction of the sample (*HS9 Study Object Sample*) from the statuette is rendered through the *HSP19 was taken from* property, which allows for the preservation of a permanent material traceability chain to the original *HS6 Study Object*. The *HS10 Sample Preparation* class is used to represent the process the sample was subjected to before analysis, specifically the chemical purification and liquid ion chromatography, defined via *HS4 Method* . The analysis for the determination of isotopic ratios is represented as a *HS3 Analysis*, using an *HS11 Scientific Device* (the thermal ionization mass spectrometer). The NIST SRM 981 reference material used for calibration operations required to ensure the validity of the measurement process is represented as a physical objects (*E19*). The analysis results (*HSP16 produced dataset*) in the generation of a *HS13 Scientific Dataset* containing the lead isotopic ratios. The geographical assessment, performed by comparing data obtained from TIMS analysis with those from European deposits (*P16 used specific object* –> *E31 Document*), is described using the *HS1 Scientific Activity* class and results (*HSP16*) in the creation of a new set of information (*HS13*) evidencing the possible geographic provenance of the lead in the statuette.

The interpretative transition from these measurements to a geographical provenance assessment is captured again through the argumentation features of CRMinf and is represented as an *I7 Belief Adoption*, structured also in this case as a set of *E12 Production* –> *P7 took place at* –> *E53 Place* propositions, which concludes a new *I2 Belief* regarding the statuette's provenance. By comparing the results with the known Aegean lead ore deposits, the authors concluded (*J2 concluded that*) that the production of the figurine (*E12 Production*) probably took place at Keos or Laurion (*E53 Place*), rather than in ancient Corinth. This formal modeling thus exposes the tension between unverified tradition and empirical data, transforming the isotopic counts into historically meaningful knowledge regarding ancient trade routes and manufacturing practices.

The ontological representation of this case study is illustrated in Figure 3.

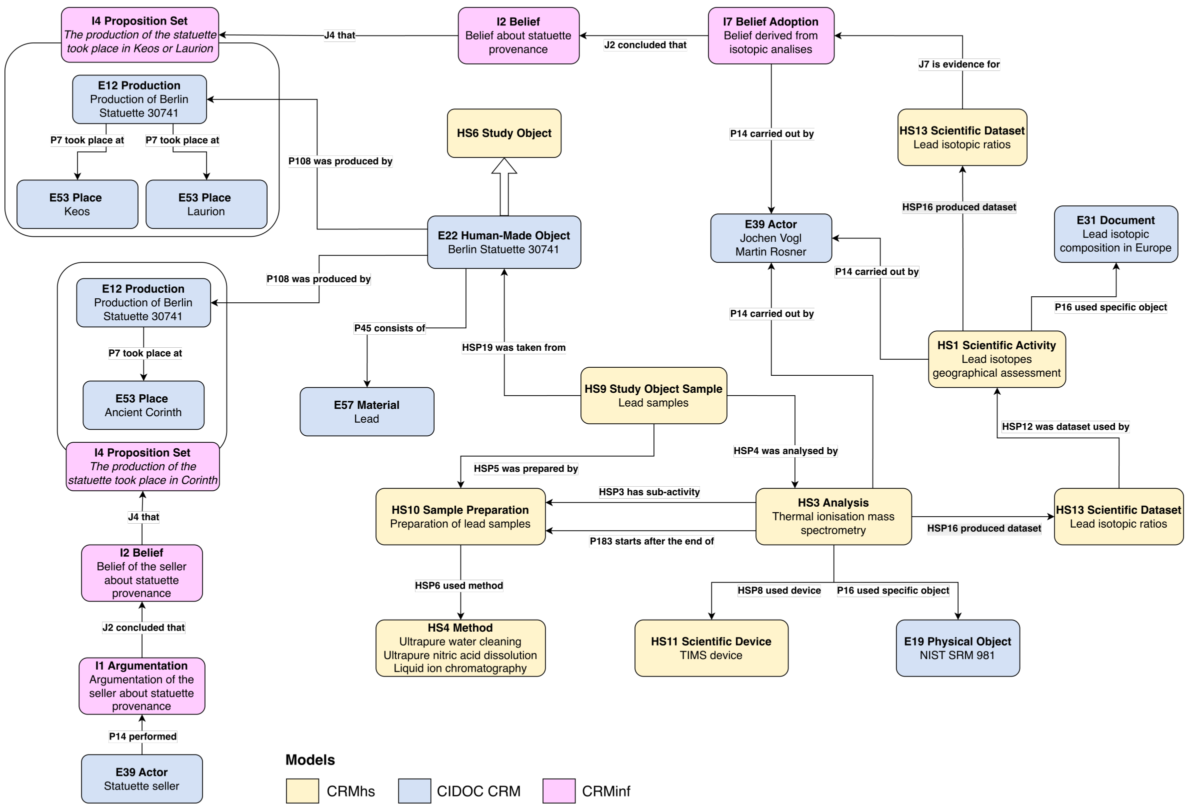

Figure 3. Semantic representation of the scientific analysis on the lead statuette 30741 from the Antikensammlung der Staatlichen Museen zu Berlin.

4.2. Text and Science: Scientific Investigations on Inscriptions Readability

The second case study examines scientific analyses designed to enhance the readability of epigraphic texts. In the study of antiquity, epigraphic texts constitute a fundamental source of direct documentation. However, their readability is frequently compromised by the poor preservation of the surfaces on which they were inscribed. This case study focuses specifically on cases involving severely weathered and abraded stone surfaces.

The application of non-invasive analytical techniques has proven essential for reconstructing text from naturally eroded surfaces, even when only faint traces of writing remain. Among these techniques, X-ray fluorescence (XRF), which we will discuss in this case study, stands out, alongside laser scanning, as demonstrated by the study of the Forum Romanum cippus. From the laser scans, multiple renderings with contour lines and a Digital Elevation Model (DEM) were generated [53].

An example of the application of XRF is documented in a 2004 study led by Cornell University in collaboration with the Cornell High-Energy Synchrotron Source [54] . This analysis investigated the elemental composition of inscribed marble surfaces, with the objective of identifying trace elements associated with ancient pigments, inscription tools, or subsequent alterations. The study exploited the high-resolution, non-invasive chemical imaging capabilities of the synchrotron.

Using a synchrotron-generated 16 keV X-ray beam, researchers performed non-destructive fluorescence mapping on inscribed marbles. The incident beam, normalised via ion chambers, was directed at the stone surface to excite specific elements. An energy-dispersive detector and multi-channel analyser (MCA) recorded the resulting spectra, with corrections applied for dead-time and surface-induced attenuation.

The setup was calibrated to detect elements ranging from calcium to strontium (K lines) and barium to bismuth (L lines), effectively mapping ancient pigments and tool fragments. Motorised stages provided high-resolution horizontal (20 µm) and vertical (1 µm) scanning, automated through SPEC software. Finally, MATLAB-based analysis produced colour-coded images of elemental intensity, demonstrating how synchrotron XRF provides crucial microscale chemical insights for cultural heritage conservation and archaeometry.

Three marble inscriptions were analysed: CIL VI 35066, CIL VI 12139, and IG II² 1969, all housed at Columbia University's Butler Library in New York (with inventory numbers NY.CU.Butl.L.13, 407, and 475, respectively). CIL VI 35066 and CIL VI 12139 are funerary inscriptions from Rome, composed in Latin. CIL VI 35066, recovered near the Basilica of Saint Paul Outside the Walls on Via Ostiense, is the epitaph of two freedmen and a freedwoman, dating to the 1st century CE (http://www.trismegistos.org/text/276270; https://usepigraphy.brown.edu/projects/usep/inscription/NY.NY.CU.Butl.L.13/). CIL VI 12139 features two epitaphs: one for the freedwoman Apollinaris, and another for three freedpeople of Aulus, dating to the first half of the 2nd century CE (http://www.trismegistos.org/text/276265; https://usepigraphy.brown.edu/projects/usep/inscription/NY.NY.CU.Butl.L.407/). IG II² 1969 is a Greek inscription listing ephebic friends from Athens, dating to 44/45 CE (https://www.atticinscriptions.com/inscription/IGII2/1969).

For the purpose of this paper, inscription CIL VI 12139, featuring a naturally eroded surface with some sections nearly completely abraded, has been chosen. The lead fluorescence imaging of this inscription successfully reveals the original lettering, facilitating the recovery of text that that had become nearly illegible. For example, in line 6, a more reliable reading of IVSAL is recovered from the visible traces on the marble. As a result, the relevant portion of the line can now be better understood as *PopillIVS A(uli) L(ibertus)*. IG II² 1969 displays shallower incisions compared to the two Latin inscriptions. With the exception of a small damaged area in the lower right corner, the inscription remains fully legible.

The study further revealed that areas subject to greater wear demonstrated reduced iron and zinc fluorescence, but elevated calcium fluorescence compared to less worn regions. The presence of copper, lead and zinc suggests that bronze-alloy tools may have been used for the engraving process. Additionally, the data suggests that iron tools were employed for initial surface preparation, while the iron detected in inscribed regions may also be attributed to iron-based red pigments.

In terms of CRMhs the marble artefact can be modelled as a *E22 Man-Made Object* found (*O19 was object found through* –> *S19 Encounter Event*) in Rome (*E53 Place*). Since *E22* is a subclass of *H6 Study Object* the abraded surfaces on the object subjected to investigation can be defined as *HS8 Study Object Area*, and identified with specific coordinates (*HSP21 has object surface coordinates* –> *E94 Space Primitive*). The object bears an inscription (*P46 is composed of*) that can be modeled using the *TX1 Written Text* class of CRMtex, another extension of the CIDOC CRM designed for encoding information from ancient texts [55–57]. *TX7 Written Text Segment* can instead be used for the illegible text segment (*TXP4 has segment*) that will be the subject of the reconstruction supported by scientific analysis.

The synchrotron-based X-ray fluorescence mapping is represented as *HS3 Analysis*, targeting the *HS8* area through the *HSP4 was analised by* property. This analytical activity involves a complex instrumental setup modeled as HS11 Scientific Device, which incorporates functional sub-units such as the beamline excitation source and the energy-dispersive detector, represented as *HS12 Device Component* and linked via *HSP9 used component*. Detailed analytical configurations, such as beam energy and spatial scanning resolution, are recorded as *HS14 Analysis Settings* through the *HSP13 used settings* property. The data processing stages (*HS1 Scientific Activity*) are governed by MCASPEC and MATLAB, both instantiated as *D14 Software* and linked to the workflow through *HSP10 used software*. The lead flurescence information resulting from this data analysis (*HS13 Scientific Dataset*), is rendered as an image (*P165 incorporates* –> *E36 Visual Item*) showing the lost letters on the inscription.

The process of reading and decoding the text through this image (*P16 was used for*) is modeled using CRMtex. The *TX5 Text Recognition* class of this model describes the recognition process (*TXP10 Deciphered Text*) of the text segment no longer visible on the inscription (*TXT7 Written Text Segment*), and allows for the documentation of the reconstruction (*TXP15 Recorded Correspondence*) of all the graphemes of the lost sequence (*TX12 Grapheme Sequence*).

The semantic representation of this case study is illustrated in Figure 4.

 doi:10.20944/preprints202603.2289.v1

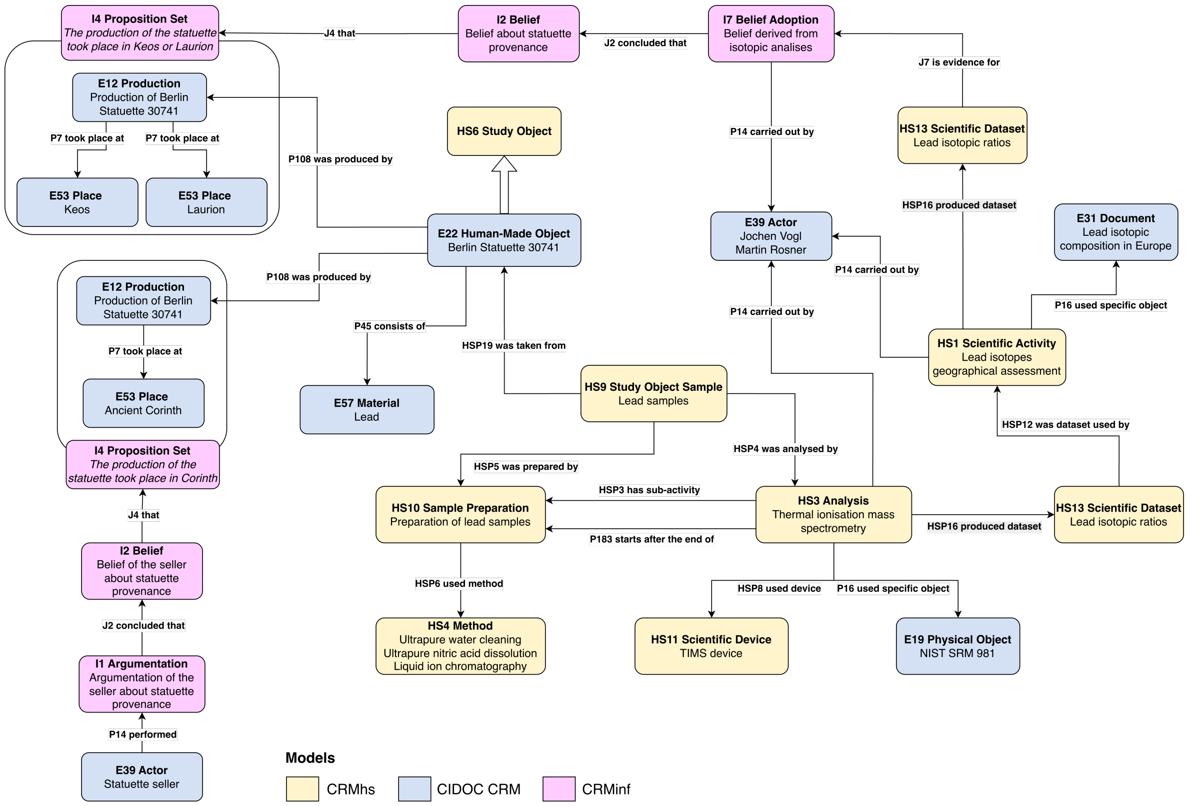

Figure 4. Semantic representation of the XRF analysis on inscription CIL VI 12139 from Rome.

5. Discussion: CRMhs as a Research Tool

The implementation of CRMhs in the two case study presented above demonstrates a significant shift from traditional, descriptive documentation toward a dynamic, evidentiary framework. By leveraging the CRMinf extension, the model successfully bridges the epistemological gap between raw analytical datasets and high-level archaeological interpretations. This is exemplified in the first example by the explicit modelling of the tension between traditional provenance claims and isotopic evidence, where CRMinf structures the transition from belief to scientifically grounded interpretation. In the second case study, focused on epigraphic readability, it is also shown as the same framework enables the integration of synchrotron-based XRF data, image processing outputs and textual reconstruction processes, demonstrating how materially grounded observations can directly inform philological interpretation. Together, these examples illustrate how CRMhs operates across different epistemic layers, linking analytical measurements, interpretative reasoning and cultural meaning within a single coherent representation.

More fundamentally, the framework reveals its full potential in its capacity to integrate scientific and cultural information within a unified semantic structure, thereby addressing the core epistemic ambition of Heritage Science itself. This integration is made possible through the precise definition of classes and properties that closely reflect the actual articulation of analytical processes, research practices, and their constituent elements. Rather than imposing an external schema, CRMhs mirrors the conceptual and operational vocabulary of the domain, enabling a faithful representation of both experimental procedures and interpretative outcomes. The alignment of CRMhs with the CIDOC CRM ecosystem is the key to guarantee that scientific data does not remain trapped within institutional silos but become immediately disposable and reusable by different disciplines and in different scenarios, an approach that addresses one of the most persistent bottlenecks in Heritage Science, namely the loss of scientific context during the transition from the laboratory to the final publication. Unlike standard metadata schemas, CRMhs treats every analytical result as the outcome of a specific, situated activity, thereby preserving the entire "chain of evidence" required for scientific reproducibility.

 doi:10.20944/preprints202603.2289.v1

Furthermore, the operational efficacy of the CRMhs framework can be significantly amplified when integrated with comprehensive, FAIR-compliant data resources such as the INFRA-ART Spectral Library. By aligning ontological precision with the empirical depth offered by similar resources it is possible to establish a robust foundation for sustainable interdisciplinary collaboration, ensuring that the material evidence required for heritage preservation is both findable and retraceable across the international research community.

Within the Heritage Digital Twin paradigm, CRMhs facilitates the convergence of disparate data layers, spanning geometric and spatial representations, analytical datasets, conservation histories, and interpretative insights, into a singular, semantically unified knowledge graph [58–60]. Consequently, the twin transcends its role as a mere visual surrogate to become a knowledge-driven system, wherein every observable feature of the asset is rigorously mapped to the underlying scientific protocols, instrumentation and datasets. Thus, the digital representation emerges as an cognitive environment rather than a static construct, fostering an architecture for simulation and predictive reasoning firmly anchored in documented evidence. This aligns with the RHDT framework's vision of reactive, adaptive twins, whose virtual states evolve in synchrony with the physical object and its broader context.

By bridging the cultural, historical, scientific and environmental dimensions, CRMhs serves as the essential semantic scaffolding for sophisticated heritage digital infrastructures and can be understood as the essential semantic link that connects and harmonises the dual nature of Heritage Science, that is, its scientific and its cultural dimensions. By making this relationship explicit and machine-interpretable, the model not only enhances data interoperability, but also supports a more reflexive and transparent form of knowledge production, in which empirical evidence and cultural interpretation are continuously aligned.

6. Conclusions

The development and implementation of CRMhs seems to move toward a resilient and adaptive framework for Heritage Science, and the preliminary results obtained are encouraging. The various applications demonstrate that the ontology could constitute a functional instrument capable of validating complex interdisciplinary hypotheses based on scientic information. However, the inherent complexity of the discipline, with its myriad of analytical facets and evolving methodologies, necessitates a sustained period of further experimentation. Heritage Science encompasses a vast spectrum of scenarios, from the micro-analysis of organic binders to the environmental monitoring of architectural complexes, each requiring specific ontological calibrations that must be tested across diverse laboratory environments.

As is characteristic of all sophisticated semantic artefacts, the CRMhs model is perceived as an evolving construct, subject to continuous modification and refinement. These iterative adjustments will not emerge from theoretical speculation but will be dictated by the practical exigencies of real-world applications. The ARTEMIS project, with its focus on the integration of reactive data within heritage digital infrastructures, constitutes the natural and primary testing ground for this evolution. By serving as the semantic backbone for high-resolution digital twins, the model will be refined to handle the dynamic flow of information required for predictive conservation and simulated restoration. This developmental path could be further enriched by the unparalleled expertise of prestigious international institutions, including the Opificio delle Pietre Dure, CNR, and the Royal Institute for Cultural Heritage (KIK-IRPA). The synergy between these laboratories and the CRMhs framework is of paramount importance, ensuring that the ontology remains deeply rooted in actual scientific practice while attaining the scalability required for global interoperability.

The ultimate objective of this research trajectory is the definition of a complete and integrated ecosystem where diverse ontological models work in perfect harmony. Within this environment, CRMhs provides the specialised technicality for Heritage Science procedures, while seamlessly interlinking with CRMsci for physical observations, CRMinf for evidentiary argumentation, and CRMtex for the philological and epigraphic analysis of ancient inscriptions. This holistic approach ensures that no piece of information exists in isolation, but rather contributes to a robust, machine-interpretable "chain of evidence" that preserves the context, provenance and uncertainty of every scientific claim.

Furthermore, the sustainability of this ecosystem depends on the development of supporting digital instruments. Future efforts will focus on creating user-centric interfaces designed specifically for research laboratories. These tools, following the trajectory set by the THESPIAN suite and AI-assisted pipelines, will simplify the encoding of complex metadata, making the formal documentation process an organic part of the scientific workflow rather than an external administrative burden. Additionally, scientific data constitutes the essential foundation for the next generation of Heritage Digital Twins, and our initial implementations in this area have already yielded significant results, proving that semantic integration is the key to transforming digital surrogates into truly cognitive systems [39,44]. This trajectory will be further expanded in upcoming research and publications where we will detail the specific methodologies for harmonizing multi-layered analytical datasets within the Heritage Digital Twin environment.

In conclusion, the significance of CRMhs extends beyond technical standardisation since it represents a fundamental commitment to the long-term preservation of our collective memory as a structured, verifiable, and evolving body of knowledge. As Heritage Science enters a more mature, data-driven era, such semantic scaffolding will be essential to ensure that the scientific biography of our cultural assets remains as resilient and accessible as the physical objects themselves. This ensures that the material evidence required for heritage protection is not only findable but remains truly retraceable for the international research community of the future.

Author Contributions: All authors have equally contributed to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original data presented in the study are openly available on the dedicated CRMhs GitHub repository at https://github.com/vastlab-dev/CRMhs. Documentation is also available at https://doi.org/10.5281/zenodo.18956182.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. House of Lords. Science and Technology Committee. Science and Heritage - 9th Report of Session 2005–06. Chapter 2: Science, technology and cultural heritage, 2006. Available online: https://publications.parliament.uk/pa/ld200506/ldselect/ldsctech/256/256.pdf (accessed on 16 March 2026).
2. Kennedy, C.; Penman, M.; Watkinson, D.; Emmerson, N.; Thickett, D.; Bosche, F.; Forster, A.; Grau-Bove, J.; Cassar, M. Beyond Heritage Science: A Review. *Heritage* **2024**. https://doi.org/10.3390/heritage7030073.
3. Pinar, G.; Sterflinger, K. Natural sciences at the service of art and cultural heritage: an interdisciplinary area in development and important challenges. *Microbial Biotechnology* **2021**, *14*, 806–809. https://doi.org/10.1111/1751-7915.13766.
4. Nanetti, A. Defining Heritage Science: A Consilience Pathway to Treasuring the Complexity of Inheritable Human Experiences through Historical Method, AI, and ML. *Complexity* **2021**, *2021*, 4703820. https://doi.org/https://doi.org/10.1155/2021/4703820.
5. Douglas-Jones, R.; Hughes, J.; Jones, S.; Yarrow, T. Science, value and material decay in the conservation of historic environments. *Journal of Cultural Heritage* **2016**, *21*, 823–833. https://doi.org/10.1016/j.culher.2016.03.007.
6. France, F. Heritage Science – Perspective, Provenance, Preservation. In Proceedings of the Archiving Conference, 2024. https://doi.org/10.2352/issn.2168-3204.2024.21.1.1.
7. euroCRIS. CERIF (Common European Research Information Format). https://eurocris.org/services/main-features-cerif, 2026. (accessed on 13 March 2026).
8. Schildhauer, M.; Jones, M.B.; Bowers, S.; Madin, J.; Krivov, S.; Pennington, D.; Villa, F.; Leinfelder, B.; Jones, C.; O'Brien, M. OBOE: the Extensible Observation Ontology, version 1.2, 2016. https://doi.org/10.5063/F1125R0F.

9. Jorg, B. CERIF: The Common European Research Information Format Model. *Data Science Journal* **2010**, *9*, CRIS24–CRIS31. https://doi.org/10.2481/dsj.CRIS4.
10. Madin, J.S.; Bowers, S.; Schildhauer, M.P.; Jones, M.B. An ontology for describing and synthesizing ecological observation data. *Ecological Informatics* **2007**, *2*, 279–296. https://doi.org/10.1016/j.ecoinf.2007.05.004.
11. Könnecke, M.; Akeroyd, F.A.; Bernstein, H.J.; Brewster, A.S.; Campbell, S.I.; Clausen, B.; Cottrell, S.; Hoffmann, J.U.; Jemian, P.R.; Männicke, D.; et al. The NeXus data format. *Journal of Applied Crystallography* **2015**, *48*, 301–305. https://doi.org/10.1107/S1600576714027575.
12. Hurt, R.; Christensen, L.L.; Gauthier, A. Astronomy Visualization Metadata (AVM), Version 1.1. International Virtual Observatory Alliance / Virtual Astronomy Multimedia Project, 2008. Available online: https://www.virtualastronomy.org/avm_metadata.php.
13. Bernstein, H.J.; Bollinger, J.C.; Brown, I.D.; Gražulis, S.; Hester, J.R.; McMahon, B.; Spadaccini, N.; Westbrook, J.D.; Westrip, S.P. Specification of the Crystallographic Information File format, version 2.0. *Journal of Applied Crystallography* **2016**, *49*, 277–284. https://doi.org/10.1107/S1600576715021871.
14. International Mineralogical Association.; RRUFF Project. IMA Mineral List with Database of Mineral Properties. Online resource, 2026. Available online: https://rruff.info/ima/ (accessed on 12 March 2026).
15. The Mineralogy Database. https://webmineral.com/ (accessed on 13 March 2026).
16. NASA STI Program. NASA Thesaurus. Online resource, 2026. Available online: https://www.sti.nasa.gov/nasa-thesaurus/ (accessed on 12 March 2026).
17. CHAMEO. CHAracterisation MEthodology Ontology. https://emmo-repo.github.io/domain-characterisation-methodology (accessed on 13 March 2026).
18. Del Nostro, P.; Goldbeck, G.; Toti, D. CHAMEO: An ontology for the harmonisation of materials characterisation methodologies. *Applied Ontology* **2022**, *17*, 1–21. https://doi.org/10.3233/AO-220271.
19. Larsson, Å.M.; Bornsäter, B.; Hacke, M. Developing practices for FAIR and linked data in Heritage Science. *npj Heritage Science* **2025**, *13*, 53. https://doi.org/10.1038/s40494-025-01598-x.
20. Cortea, I.M. Towards FAIR Data Management in Heritage Science Research: Updates and Progress on the INFRA-ART Spectral Library. *Heritage* **2024**, *7*, 2569–2585. https://doi.org/10.3390/heritage7050123.
21. E-RIHS. European Research Infrastructure for Heritage Science. https://www.e-rihs.eu/ (accessed on 13 March 2026).
22. H2IOSC Consortium. Humanities and Cultural Heritage Italian Open Science Cloud (H2IOSC). https://www.h2iosc.cnr.it/ (accessed on 13 March 2026).
23. Scarpa, E.; Valente, R.; Rossi, I. Modeling an Ontology for Heritage Science: Challenges and Key Strategies. In Proceedings of the Convegno Annuale AIUCD2025. Diversità, Equità e Inclusione: Sfide e Opportunità per l'Informatica Umanistica nell'Era dell'Intelligenza Artificiale; Rebora, S.; Rospocher, M.; Bazzaco, S., Eds., Verona, 2025; Quaderni di Umanistica Digitale, pp. 499–504. Available online: https://amsacta.unibo.it/id/eprint/8380/(accessed on 13 March 2026).
24. ICOM/CIDOC Documentation Standards Group. The CIDOC Conceptual Reference Model (ISO 21127). https://www.cidoc-crm.org/. (accessed on 13 March 2026).
25. Definition of the CRMsci. An Extension of CIDOC CRM to support scientific observation. Available online: https://cidoc-crm.org/crmsci (accessed on 12 March 2026).
26. The Getty Research Institute. Art & Architecture Thesaurus (AAT). https://www.getty.edu/research/tools/vocabularies/aat/ (accessed on 13 March 2026).
27. PeriodO. A gazetteer of periods for linking and visualizing data. https://perio.do/ (accessed on 13 March 2026).
28. PLEIADES. A community-built gazetteer and graph of ancient places. https://pleiades.stoa.org/ (accessed on 13 March 2026).
29. Castelli, L.; Felicetti, A. Modelli e strumenti semantici per la documentazione dell'Heritage Science. In Proceedings of the Data Revolution. Conferenza GARR_18 (Cagliari, 3-5 ottobre 2018). Selected papers. GARR, 2018, pp. 56–59. https://doi.org/10.26314/GARR-Conf18-proceedings-10.
30. Niccolucci, F.; Meghini, C.; Felicetti, A.; Pezzati, L. Condivisione dei dati sui beni culturali: DIGILAB, l'esperienza di ARIADNE e di E-RIHS. In Proceedings of the The data way to Science. Conferenza GARR_17 (Venezia, 15-17 novembre 2017). Selected papers. GARR, 2017, pp. 129–132. https://doi.org/10.26314/GARR-Conf17-proceedings-24.
31. Niccolucci, F.; Felicetti, A. A CIDOC CRM-based Model for the Documentation of Heritage Sciences. In Proceedings of the 2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held jointly

with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018), 2018, pp. 1–6. https://doi.org/10.1109/DigitalHeritage.2018.8810109.

32. Castelli, L.; Felicetti, A.; Proietti, F. Heritage Science and Cultural Heritage: Standards and Tools for Establishing Cross-Domain Data Interoperability. *International Journal on Digital Libraries* **2021**, *22*, 279–287. https://doi.org/10.1007/s00799-019-00275-2.
33. Bombini, A.; Castelli, L.; Dell'Agnello, L.; Felicetti, A.; Giacomini, F.; Niccolucci, F.; Taccetti, F. CHNet Cloud: an EOSC-based Cloud for Physical Technologies Applied to Cultural Heritages. In Proceedings of the Sostenibile/Digitale. Conferenza GARR_21 (Online, 7-16 giugno 2021). Selected papers. GARR, 2021, pp. 41–46. https://doi.org/10.26314/GARR-Conf21-proceedings-09.
34. Pezzati, L.; Felicetti, A. DIGILAB: A New Infrastructure for Heritage Science. *ERCIM News* **2017**, *111*. Available online: https://ercim-news.ercim.eu/en111/special/digilab-a-new-infrastructure-for-heritage-science (accessed on 12 March 2026).
35. ARIADNE. Advanced Research Infrastructure for Archaeological Dataset Networking in Europe. https://www.ariadne-research-infrastructure.eu/, 2026. (accessed on 13 March 2026).
36. 4CH. European Competence Centre for the Conservation of Cultural Heritage. https://www.4ch-project.eu/, 2023. (accessed on 13 March 2026).
37. ARTEMIS. Applying Reactive Twins to Enhance Monument Information Systems. https://www.artemis-twin.eu/, 2026. (accessed on 13 March 2026).
38. Middle, S.; Richards, J.; Constantopoulos, P.; Costantini, A.; Dritsou, V.; Felicetti, A.; Raffaelli, M.; Somenzi, M. D7.1 Infrastructure Design and Setup. Project deliverable, ARTEMIS Project (EU Horizon Europe grant 101188009), 2025. https://doi.org/10.5281/zenodo.17142049.
39. Felicetti, A.; Himmiche, A.; Somenzi, M. Knowledge Graphs and Artificial Intelligence for the Implementation of Cognitive Heritage Digital Twins. *Applied Sciences* **2025**, *15*. https://doi.org/10.3390/app151810061.
40. Bordalo, R.; Bottaini, C.; Candeias, A. A Framework Design for Information Management in Heritage Science Laboratories. *J. Comput. Cult. Herit.* **2021**, *14*. https://doi.org/10.1145/3417304.
41. Felicetti, A.; Niccolucci, F. CRMhs Ontology - A CIDOC CRM extension for the semantic modelling of Heritage Science information, 2026. https://doi.org/10.5281/zenodo.18956182.
42. Felicetti, A.; Niccolucci, F. CRMhs: A semantic model for Heritage Science, 2026. https://github.com/vastlab-dev/CRMhs (accessed on 13 March 2026).
43. Mazzinghi, A.; Ruberto, C.; Giuntini, L.; Mandò, P.A.; Taccetti, F.; Castelli, L. Mapping with Macro X-ray Fluorescence Scanning of Raffaello's Portrait of Leo X. *Heritage* **2022**, *5*, 3993–4005. https://doi.org/10.3390/heritage5040205.
44. Felicetti, A.; Niccolucci, F. Artificial Intelligence and Ontologies for the Management of Heritage Digital Twins Data. *Data* **2025**, *10*. https://doi.org/10.3390/data10010001.
45. Versnel, H.S. Prayer and Curse. In *The Oxford Handbook of Ancient Greek Religion*; Eidinow, E.; Kindt, J., Eds.; Oxford University Press: Oxford, 2015; pp. 447–462. https://doi.org/10.1093/oxfordhb/9780199642038.013.31.
46. Eidinow, E. *Oracles, Curses, and Risk Among the Ancient Greeks*; Oxford University Press: Oxford, 2007.
47. Jordan, D.R. Defixiones from a well near the southwest corner of the Athenian Agora. *Hesperia* **1985**, *54*, 205–255. Available online: https://www.ascsa.edu.gr/uploads/media/hesperia/147887.pdf (accessed on 25 March 2026).
48. Gordon, R.L. Showing the Gods the Way: Curse-tablets as Deictic Persuasion. *Religion in the Roman Empire* **2015**, *1*, 148–180. https://doi.org/10.1628/219944615X14296073073539.
49. Curbera, J.B.; Giannobile, S. A 'Voodoo Doll' from Keos in Berlin's Antikensammlung. In *The Materiality of Magic*; Boschung, D.; Bremmer, J.N., Eds.; Fink: Paderborn, 2015; pp. 123–126.
50. Curbera, J.B. From the Magician's Workshop: Notes on the Materiality of Greek Curse Tablets. In *The Materiality of Magic*; Boschung, D.; Bremmer, J.N., Eds.; Fink: Paderborn, 2015; pp. 97–122.
51. Vogl, J.; Rosner, M. Appendix: Lead Isotope Analysis of an Ancient Voodoo Doll. In *The Materiality of Magic*; Boschung, D.; Bremmer, J.N., Eds.; Fink: Paderborn, 2015; pp. 127–132.
52. Definition of the CRMinf. An Extension of CIDOC CRM for integrating metadata about argumentation and inference making in descriptive and empirical sciences. Available online: https://cidoc-crm.org/crminf (accessed on 12 March 2026).
53. Rocca, G.; Sarullo, G. *Il Cippo del Foro: nuove letture e prospettive euristiche*; De Gruyter: Berlin Boston, 2025.

54. Powers, J.; Dimitrova, N.; Huang, R.; Smilgies, D.M.; Bilderback, D.H.; Clinton, K.; Thorne, R.E. X-ray Fluorescence Recovers Writing from Ancient Inscriptions. *Zeitschrift für Papyrologie und Epigraphik* **2005**, *152*, 221–227.
55. Definition of the CRMtex. An Extension of CIDOC CRM for the study of ancient texts. Available online: https://cidoc-crm.org/crmtex (accessed on 12 March 2026).
56. Felicetti, A.; Murano, F. Scripta manent: a CIDOC CRM semiotic reading of ancient texts. *International Journal on Digital Libraries* **2017**, *18*, 263–270. https://doi.org/10.1007/s00799-016-0189-z.
57. Felicetti, A.; Murano, F. Ce qui est écrit et ce qui est parlé. CRMtex for modelling textual entities on the Semantic Web. *Semantic Web* **2021**, *12*, 169–180. https://doi.org/10.3233/SW-200418.
58. Niccolucci, F.; Felicetti, A.; Hermon, S. Populating the Data Space for Cultural Heritage with Heritage Digital Twins. *Data* **2022**, *7*. https://doi.org/10.3390/data7080105.
59. Niccolucci, F.; Felicetti, A. Digital Twin Sensors in Cultural Heritage Ontology Applications. *Sensors* **2024**, *24*. https://doi.org/10.3390/s24123978.
60. Niccolucci, F.; Felicetti, A. Digital Twins and the Stendhal Syndrome. *Computers* **2025**, *14*. https://doi.org/10.3390/computers14040136.