NATIONAL LEADERSHIP GRANTS FOR LIBRARIES

Sample Application LG-260266-OLS-26
Project Category: Applied Research

Indiana University

Amount awarded by IMLS:	$268,255
Amount of cost share:	$0

Indiana University will investigate how the CIDOC Conceptual Reference Model can support research-oriented knowledge structures capable of integrating these heterogeneous forms of cultural heritage data. CIDOC-Conceptual Reference Model was developed to enable the exchange and integration of cultural heritage information across institutions and systems. While widely used for museum and cultural heritage documentation and metadata interoperability, this project seeks for the CIDOC-Conceptual Reference Model to be applied to modeling the research processes, analytical measurements, and evolving interpretations through which knowledge about cultural heritage objects is produced. The project will investigate how CIDOC-Conceptual Reference Model based modeling strategies can structure and integrate complex cultural heritage research data, including digital representations, chemical and spectrometric analyses, computational results, and scholarly interpretation. The project will use the Marble Map of Rome Project as a pilot dataset.

Attached are the following components excerpted from the original application.

- Narrative
- Schedule of completion

When preparing an application for the next deadline, be sure to follow the instructions in the most recent Notice of Funding Opportunity for the grant program to which you are applying.

INTRODUCTION

Cultural heritage research is generating increasingly complex forms of data that challenge the descriptive capacities of traditional library metadata systems. High-resolution three-dimensional scanning, chemical and spectrometric analysis, advanced imaging techniques, and computational methods now produce datasets that capture physical, material, and analytical properties of artifacts in unprecedented detail. Libraries and archives are increasingly responsible for stewarding these datasets alongside traditional records and digital surrogates. However, existing descriptive frameworks remain largely optimized for cataloging textual records and two-dimensional images, limiting their ability to represent relationships among physical objects, digital representations, scientific measurements, and evolving scholarly interpretations. At the same time, emerging research methods increasingly generate computational features and analytical datasets derived from digital representations, raising new questions about how such evidence can be incorporated into cultural heritage knowledge structures.

This project investigates how the CIDOC Conceptual Reference Model (CIDOC-CRM) can support research-oriented knowledge structures capable of integrating these heterogeneous forms of cultural heritage data. CIDOC-CRM was developed to enable the exchange and integration of cultural heritage information across institutions and systems. While widely used for museum and cultural heritage documentation and metadata interoperability, CIDOC-CRM has been less frequently applied to modeling the research processes, analytical measurements, and evolving interpretations through which knowledge about cultural heritage objects is produced.

Indiana University proposes a two-year Applied Research project that investigates how CIDOC-CRM–based modeling strategies can structure and integrate complex cultural heritage research data, including digital representations, chemical and spectrometric analyses, computational results, and scholarly interpretation. The project will use the Marble Map of Rome Project as a pilot dataset. This corpus includes hundreds of marble fragments documented through high-resolution three-dimensional scanning as well as associated analytical datasets and extensive scholarly literature. By modeling the relationships among fragments, digital representations, scientific measurements, and scholarly interpretations, the project will explore how CIDOC-CRM can support modular knowledge structures designed for research rather than solely for cataloging and discovery.

The project will also investigate methods for integrating information extracted from the scholarly literature associated with the Marble Map corpus. Using experimental AI-assisted workflows, the project will evaluate how references to fragments, images, and catalog identifiers appearing in journal articles and archaeological publications can be identified and incorporated into the CIDOC-based data model. These experiments will test whether a research-oriented knowledge structure can connect physical artifacts, analytical datasets, computational results, and scholarly discourse within a unified semantic framework.

The outcomes of this project will include prototype implementations, modeling guidelines, and recommendations for libraries and archives managing complex research datasets. By exploring how CIDOC-CRM can support the integration of scientific analysis, digital representations, and scholarly knowledge, the project will contribute new approaches for representing research-grade cultural heritage data within library information infrastructures.

PROJECT JUSTIFICATION

Transformation of Cultural Heritage Research Data

In archaeology, art history, conservation science, and related fields, digital recording technologies and analytical methods are generating new categories of research data. High-resolution three-dimensional scanning, advanced imaging techniques, chemical and spectrometric analysis, and computational modeling now produce datasets capable of documenting the physical and material characteristics of artifacts with unprecedented precision. These methods enable forms of analysis—such as quantitative comparison of surfaces, measurement of carving techniques, and computational evaluation of fragment joins—that were previously difficult or impossible to perform.

Libraries and archives are increasingly responsible for stewarding these emerging forms of research data. Digital repositories now host three-dimensional models, analytical datasets, and associated documentation alongside traditional digital collections. However, the descriptive frameworks used to organize these materials were largely developed to support cataloging and discovery of textual records and two-dimensional images. As a result, many research datasets remain stored separately from object records, described inconsistently, or disconnected from the scholarly literature and analytical processes through which knowledge about the objects is produced.

This fragmentation presents a growing challenge for libraries seeking to support digital scholarship and data-intensive humanities research. When digital representations, scientific measurements, and interpretive scholarship

cannot be connected within shared knowledge structures, the ability of researchers to integrate evidence across datasets is limited. Libraries therefore face an important infrastructure question: how can cultural heritage data models represent the complex relationships among physical objects, digital representations, analytical measurements, computational analysis, and scholarly interpretation?

CIDOC-CRM as a Conceptual Framework

The CIDOC Conceptual Reference Model (CIDOC-CRM) provides a widely adopted conceptual framework for integrating cultural heritage information across institutions and systems. Developed by the International Committee for Documentation of the International Council of Museums, CIDOC-CRM defines a formal ontology designed to support the exchange and integration of heterogeneous cultural heritage data. Its central goal is to enable disparate information sources to be connected within coherent semantic structures that allow researchers to reconstruct and interpret the past using diverse forms of evidence.

CIDOC-CRM has been widely used in museum documentation, digital heritage initiatives, and linked data projects to support interoperability between collection management systems. However, most implementations focus primarily on object description, catalog metadata, and institutional documentation. Far fewer projects have explored how CIDOC-CRM might support the representation of research processes and analytical data generated through scientific investigation and computational analysis.

This project investigates whether CIDOC-CRM can support knowledge structures capable of integrating multiple categories of cultural heritage research data. These include digital representations produced through high-resolution three-dimensional scanning, analytical measurements derived from chemical or spectrometric analysis, computational outputs generated through machine learning workflows, and interpretive relationships documented in scholarly literature. This project will evaluate how CIDOC-CRM can support a modular modeling approach in which a core representation of cultural heritage objects is linked to separate components representing digital representations, analytical measurements, computational results, and scholarly documentation.

Research Questions

This project investigates how cultural heritage knowledge structures can evolve to accommodate the increasingly complex datasets generated through contemporary research across the humanities, social sciences, and cultural heritage fields. Libraries and archives are increasingly responsible for stewarding datasets that combine physical artifacts, digital representations, analytical measurements, and scholarly interpretation. While CIDOC-CRM has become a widely adopted framework for cultural heritage interoperability, its potential for representing research environments that integrate these heterogeneous forms of evidence remains comparatively underexplored.

The project therefore addresses the following research questions:

1. How can CIDOC-CRM support the representation of research-oriented cultural heritage datasets that integrate physical objects, digital representations, analytical measurements, and scholarly interpretation within library information infrastructures?
2. How can knowledge structures represent the historical layering of scholarly documentation associated with cultural heritage collections, including successive cataloging systems, visual representations, and evolving research interpretations?
3. What modeling approaches allow analytical and computational research data—such as scientific measurements, imaging datasets, and machine-assisted analysis—to be integrated into CIDOC-based knowledge structures while maintaining interoperability with existing cultural heritage standards?
4. How can information embedded in scholarly publications—including references to objects, images, and catalog identifiers—be incorporated into cultural heritage data infrastructures through AI-assisted methods?

Using the Great Marble Map of Rome as a pilot corpus, the project will evaluate these questions through a series of prototype implementations and experimental workflows designed to test how CIDOC-based models can represent complex research environments in which artifacts, digital representations, analytical datasets, and scholarly discourse intersect.

Fragmented Artifacts and the Complexity of Cultural Heritage Data

Fragmented artifacts present particularly demanding challenges for cultural heritage data modeling. In many

archaeological and art historical contexts, objects survive only as partial remains that must be interpreted through relationships among fragments, historical documentation, and analytical measurements. These relationships are often complex and evolving, as new research produces revised interpretations and new connections among fragments.

The Great Marble Map of Rome (*Forma Urbis Romae*) represents one of the most significant examples of such a fragmented corpus. Originally carved in marble and displayed in early third-century Rome, the map depicted the architectural plan of the city in extraordinary detail. Over the centuries the monument was broken apart, leaving hundreds of fragments that survive today in museums and archaeological collections. These fragments have long been the subject of scholarly study, as they provide unique evidence for the urban layout of ancient Rome.

Recent digitization efforts have produced high-resolution three-dimensional scans of the surviving fragments, capturing geometric detail and surface characteristics at a scale suitable for computational analysis. These digital models allow scholars to examine carving techniques, analyze fracture surfaces, and test potential joins between fragments using quantitative methods. At the same time, the fragments are associated with a wide range of additional datasets, including analytical imaging, material analysis, historical catalog records, and extensive scholarly literature documenting earlier interpretations.

The Marble Map therefore presents a highly complex data environment in which physical fragments, digital representations, analytical measurements, and scholarly interpretations intersect. This complexity makes the corpus an ideal testbed for evaluating how CIDOC-CRM can support research-oriented knowledge structures capable of integrating heterogeneous cultural heritage data.

Advances in Digital Documentation of the Marble Map

Digital documentation of the Great Marble Map of Rome has developed over several decades through successive technological phases. The most influential early effort to digitize the fragments was the Stanford Digital Forma Urbis Romae Project, which created three-dimensional models of the surviving fragments using laser scanning and digital photography. That project produced the first comprehensive digital dataset for the corpus and demonstrated the potential of computational methods for analyzing fragment geometry and identifying possible joins between fragments.

While groundbreaking at the time, the Stanford dataset was produced using technologies that limited the resolution of the resulting models. The scans captured overall fragment geometry but lacked the level of surface detail necessary for certain types of analytical research. As digital recording technologies advanced, it became clear that higher-resolution models would be required to study fine features such as incision depth, carving technique, fracture morphology, and other subtle surface characteristics that may provide evidence for fragment joins or differences in carving practice.

Beginning in 2018, the Great Marble Map of Rome Project initiated a new phase of digital documentation designed to address these limitations. The project first produced a photogrammetric survey of the original display wall in the Roman Forum, recording the positions and dimensions of the clamp holes that once held the marble slabs in place. This survey created a detailed digital model of the architectural context in which the map was originally displayed.

In 2019, a team based at Indiana University conducted a major scanning campaign in Rome to create a new generation of digital models of the surviving fragments. Using handheld structured-light scanners capable of extremely high geometric precision, the team produced high-resolution three-dimensional scans of all known incised fragments of the Marble Map. These scanners capture surface geometry by projecting structured light across the object and reconstructing the resulting geometry through high-frequency imaging. The resulting models achieve accuracies approaching hundredths of a millimeter while simultaneously capturing high-resolution color texture from thousands of photographs per second.

These scans represent a substantial improvement over earlier digital models and allow scholars to analyze the fragments at a level of detail not previously possible. Features such as the slope and depth of incised lines, subtle tool marks, surface weathering patterns, and fracture geometries can now be measured quantitatively across large numbers of fragments. The availability of these high-resolution models also enables new computational approaches to the study of the monument, including machine-assisted comparison of fragment surfaces and the analysis of carving techniques across the corpus. These forms of analysis generate new categories of structured research data—such as geometric descriptors, surface metrics, and algorithmically derived indicators—that must be represented alongside traditional catalog metadata within cultural heritage knowledge structures.

At the same time, the creation of these new digital datasets introduces new challenges for cultural heritage data infrastructure. The Marble Map corpus now includes multiple generations of digital documentation: early laser-scan models, newer high-resolution structured-light scans, photogrammetric models of the display wall, and associated analytical datasets derived from these digital representations. These datasets must be connected not only to the physical fragments but also to historical catalog records, earlier scholarly reconstructions, and the extensive literature documenting the monument.

As a result, the Marble Map corpus now functions as a complex research environment in which physical artifacts, multiple generations of digital models, analytical measurements, and scholarly interpretations intersect. This environment provides an ideal test case for evaluating how CIDOC-CRM–based knowledge structures can represent evolving research datasets while preserving the historical record of how the monument has been documented and interpreted over time.

The Marble Map as a Test Case for Cultural Heritage Data Modeling

The Marble Map of Rome (*Forma Urbis Romae*) provides an unusually revealing case through which to examine the capabilities and limitations of existing cultural heritage data models. Few archaeological corpora combine such a wide range of data types and scholarly documentation. The surviving fragments of the map are associated with physical artifacts, historical catalog records, engraved reconstructions, archaeological drawings, photographic documentation, high-resolution three-dimensional scans, analytical measurements, and an extensive body of scholarly literature spanning more than three centuries.

This combination of heterogeneous evidence exposes the complexity of modeling cultural heritage research environments. The fragments themselves are archaeological objects, but they are also the subjects of long-standing scholarly interpretation and reconstruction. Over time they have been cataloged under multiple numbering systems, represented in different visual media, and incorporated into successive hypotheses about the original structure of the monument. The corpus therefore exists not simply as a collection of objects but as an evolving research assemblage in which artifacts, representations, analytical data, and scholarly interpretations intersect.

Because the Marble Map corpus contains such a rich and varied documentation history, it provides an ideal experimental environment for evaluating how cultural heritage data models can represent research processes as well as objects. Examining how these diverse forms of evidence can be connected within CIDOC-CRM–based knowledge structures allows the project to test whether existing modeling approaches can accommodate the layered and evolving nature of archaeological research datasets.

Limitations of Existing CIDOC-CRM Extensions for Research Data

CIDOC-CRM provides a widely adopted framework for representing cultural heritage information and has been extended through a number of domain-specific modules designed to address particular categories of documentation. These extensions include models for archaeological excavation processes, scientific observation, and digital provenance, among others. Together they provide an important foundation for representing cultural heritage collections and the activities associated with their documentation.

However, these extensions were largely developed to support specific documentation workflows rather than the integrated representation of complex research environments. Archaeological extensions focus primarily on modeling excavation contexts and stratigraphic relationships, while digitization models emphasize the provenance of digital objects produced through imaging and scanning processes. Other extensions address particular forms of scientific observation or measurement. While these models provide valuable tools for describing specific aspects of cultural heritage documentation, they do not fully address the modeling challenges posed by research datasets that combine physical artifacts, digital representations, analytical measurements, computational analysis, and evolving scholarly interpretation within a single knowledge structure.

The Marble Map corpus illustrates this problem clearly. The fragments of the map are associated with a wide range of data types, including high-resolution three-dimensional scans, chemical and spectrometric analyses, computational models used to test fragment joins, historical catalog records, and extensive scholarly literature documenting earlier reconstructions. These datasets are interconnected in ways that reflect not only the physical characteristics of the fragments but also the research processes through which knowledge about the monument has been produced. Modeling these relationships requires a framework capable of representing objects, analytical observations, computational results, and interpretive hypotheses while also preserving the historical layers of documentation through

which the corpus has been assembled and reinterpreted.

A further challenge emerges when considering the types of computational and analytical features generated through high-resolution digital recording and analysis. Research datasets derived from three-dimensional models increasingly include geometric descriptors, surface metrics, and algorithmically derived indicators that support interpretive reasoning about fragment relationships. In the Marble Map corpus, for example, high-resolution scans allow the extraction of features such as fracture geometry, thickness variation, surface continuity, and computational scoring models that evaluate multiple forms of evidence when proposing potential joins between fragments. Such data represent not simply empirical observations but structured representations of analytical reasoning derived from digital models. Existing CIDOC-CRM extensions provide mechanisms for describing objects, observations, and digitization processes, but they offer limited guidance for representing these kinds of derived computational features and multi-evidence analytical workflows within a unified knowledge structure. By grounding its investigation in the Marble Map dataset, the project evaluates how such features can be modeled within CIDOC-based frameworks while maintaining interoperability with existing cultural heritage standards.

Existing CIDOC-CRM extensions provide important conceptual tools for representing portions of these relationships, but they do not fully address the need to integrate heterogeneous research data and scholarly interpretation within a unified knowledge structure. In particular, the current modules do not provide a clear framework for representing computational analysis workflows, linking analytical measurements to features derived from high-resolution three-dimensional models, or integrating information extracted from the scholarly literature that documents the evolving interpretation of the fragments.

Rather than proposing an entirely new ontology, this project investigates how CIDOC-CRM can support a modular approach in which a core representation of cultural heritage objects is extended through linked modules that address the specific requirements of research-oriented datasets. The goal of the project is therefore not to replace existing CIDOC extensions, but to evaluate how the framework can be adapted to represent complex research environments in which physical artifacts, digital representations, analytical measurements, and scholarly interpretations interact.

By testing these modeling approaches using the Marble Map corpus, the project will contribute new insights into how CIDOC-CRM can support cultural heritage data infrastructures designed for research as well as documentation.

The Historical Layering of Scholarly Documentation

An additional challenge arises from the fact that cultural heritage datasets themselves possess complex histories of documentation and interpretation. The fragments of the Great Marble Map of Rome have been cataloged and reinterpreted repeatedly since the eighteenth century. Over this period the fragments have been represented in multiple forms, including engraved reconstructions, antiquarian drawings, archaeological publications, photographic documentation, and, more recently, digital models derived from high-resolution three-dimensional scanning.

Successive cataloging efforts have assigned different identifiers to fragments, sometimes introducing errors, duplications, or missing references as fragments were rediscovered, reclassified, or reassigned to new hypothetical positions within the map. Reconstructions of the map have likewise evolved over time, with scholars proposing different arrangements of fragments and different interpretations of the buildings represented on the marble surface.

As a result, the corpus exists not simply as a collection of archaeological objects but as a layered scholarly assemblage shaped by centuries of documentation, reconstruction, and interpretation. In effect, the Marble Map constitutes a palimpsest of scholarly representations, in which fragments have been repeatedly documented, interpreted, and visually reconstructed across multiple generations of research. These representations include textual descriptions, engraved images, archaeological drawings, photographic records, and digital models, each reflecting different scholarly assumptions and methodological approaches.

An important component of this project is therefore to explore how CIDOC-based knowledge structures can represent not only the fragments themselves but also the historical processes through which the corpus has been assembled, cataloged, reconstructed, and represented over time. Linking fragments to their historical visual, textual, and digital representations will allow scholars to trace the evolution of interpretations and to understand the scholarly processes through which knowledge about the monument has been constructed.

Integrating Scientific and Computational Research

In addition to digital representations, the project will examine how scientific research data can be incorporated

into CIDOC-based knowledge structures. Analytical methods such as chemical analysis, spectrometry, and imaging techniques increasingly play important roles in the study of cultural heritage objects. These methods generate measurements and observations that contribute directly to scholarly interpretation but are often stored separately from the descriptive records that libraries and archives maintain.

Representing these analytical datasets within shared semantic structures raises important modeling questions. Measurements may be associated with particular fragments, specific surfaces, or features observed through digital models. Analytical results may also contribute to interpretive hypotheses about fragment joins, carving techniques, or material provenance. Modeling these relationships requires approaches capable of representing both empirical observations and the research processes through which interpretations are produced.

The project will therefore explore how CIDOC-CRM can accommodate these research workflows while maintaining interoperability with existing cultural heritage data models. By testing the integration of analytical measurements and computational outputs within a CIDOC-based framework, the project will contribute new insights into how libraries can represent research-grade cultural heritage data.

Scholarly Literature as Data

Another important component of the project involves the integration of information derived from the scholarly literature associated with the Marble Map corpus. Over the past several centuries, archaeological publications, journal articles, excavation reports, and catalogs have documented fragments of the map and proposed interpretations of their relationships. These publications contain a large body of knowledge about the fragments, including images, fragment identifications, references to earlier cataloging systems, and descriptions of architectural features represented on the marble surface.

Much of this knowledge, however, exists only in the form of natural-language descriptions and embedded visual documentation rather than in structured data systems. As a result, valuable information about fragment identifications, historical reconstructions, and interpretive hypotheses often remains difficult to connect systematically with the digital representations and analytical datasets associated with the fragments.

The project will therefore investigate experimental methods for transforming natural-language scholarly descriptions into structured data representations that can be incorporated into CIDOC-based knowledge structures. Using AI-assisted approaches to text analysis and image recognition, the project will explore how references to fragments, images, and catalog identifiers appearing in scholarly publications can be identified and linked to corresponding entities within the knowledge structure.

These experiments will not seek to replace traditional scholarship, but rather to explore how narrative scholarly discourse can be connected to structured cultural heritage data infrastructures. By linking textual scholarship to structured representations of artifacts, digital models, and analytical datasets, the project will examine how libraries can incorporate the intellectual history of research itself into interoperable data environments.

Significance for Libraries

Libraries and archives increasingly serve as stewards of complex research datasets produced through digital scholarship and computational humanities projects. As these datasets grow in scope and complexity, the need for knowledge structures capable of integrating multiple forms of cultural heritage data becomes more pressing. By investigating how CIDOC-CRM can support research-oriented data models that connect digital representations, scientific measurements, computational analysis, and scholarly interpretation, this project addresses a critical challenge facing cultural heritage information infrastructures. The findings will provide libraries with conceptual frameworks, prototype implementations, and practical guidance for representing complex cultural heritage research data. These outcomes will help libraries support emerging forms of scholarship that depend on the integration of heterogeneous datasets and analytical methods, strengthening the role of libraries as central partners in digital humanities research.

PROJECT WORK PLAN

The project will be conducted over a two-year period through four interrelated research themes that structure the investigation of the project's research questions. While these themes provide an organizational framework for the work, many of the activities described below will proceed concurrently and will develop iteratively throughout the project. Conceptual modeling, data integration, analytical experimentation, and scholarly consultation will inform one another as the project progresses.

These research themes correspond to the project's methodological questions concerning how cultural heritage knowledge structures can integrate physical artifacts, digital representations, analytical measurements, computational analysis, and scholarly interpretation. The Great Marble Map of Rome corpus will serve as the pilot dataset through which these questions are explored.

Each research theme addresses one or more of the project's research questions. Theme 1 focuses on the development of conceptual modeling strategies using CIDOC-CRM. Theme 2 investigates how heterogeneous datasets and historical cataloging systems can be integrated into shared knowledge structures. Theme 3 explores the incorporation of analytical and computational research data, including machine-assisted methods for transforming scholarly literature into structured data representations. Theme 4 evaluates these approaches through consultation with external collaborators and disseminates the resulting frameworks to library and cultural heritage communities.

The project will be led by the Marble Map of Rome Project, based at the IU Indianapolis Arts and Humanities Institute, in collaboration with the Center for Digital Technology at the IU Indianapolis Library and Dockive, a digital archiving and database development firm specializing in cultural heritage collections incorporating three-dimensional data. The project team brings together expertise in archaeology, history, library science, digital humanities, digital visualization, artificial intelligence, and computational analysis.

Research Theme 1: Conceptual Modeling and Research Design
Primary focus: Months 1–6, continuing throughout the project

The first research theme focuses on developing the conceptual framework for representing complex cultural heritage research data using CIDOC-CRM–based knowledge structures. During this stage the research team will conduct a systematic analysis of the different categories of data associated with the Marble Map corpus. These include physical fragments, high-resolution three-dimensional scans, analytical datasets, historical catalog records, visual documentation, and scholarly publications. This theme directly addresses the project's first research question concerning how CIDOC-CRM can support research-oriented knowledge structures that integrate diverse cultural heritage datasets.

A central objective of this theme is to identify how these diverse forms of evidence can be represented within CIDOC-CRM while preserving the historical and scholarly relationships associated with the fragments. Particular attention will be given to modeling the historical layers of documentation and reconstruction that characterize the Marble Map corpus. Because fragments have been cataloged and interpreted repeatedly since the eighteenth century, the dataset includes multiple identifier systems, historical representations, and competing scholarly interpretations.

The research team will develop initial conceptual models describing how fragments, digital representations, analytical data, and scholarly references may be represented within a unified knowledge structure. These models will be refined through ongoing collaboration among project participants and through consultation with scholars working on related cultural heritage data initiatives.

Outputs of this theme will include:

- conceptual modeling strategies for representing Marble Map data within CIDOC-CRM
- documentation of the data categories and relationships present in the corpus
- initial prototype schema structures supporting the integration of historical identifier systems and documentation layers

These conceptual frameworks will guide the prototype implementations developed through the subsequent research themes.

Research Theme 2: Data Integration and Prototype Development
Primary focus: Months 4–12

The second research theme focuses on implementing prototype knowledge structures that integrate the various datasets associated with the Marble Map corpus. This work will begin during the latter stages of Theme 1 and will continue through the first year of the project. This work addresses the project's second research question by examining how cultural heritage knowledge structures can represent the historical layering of cataloging systems, visual representations, and scholarly interpretations associated with complex research corpora.

The prototypes will integrate fragment metadata, high-resolution three-dimensional scans, historical documentation, and selected analytical datasets. The high-resolution scans of the fragments provide a particularly valuable dataset for this work because they capture geometric and surface characteristics with sufficient precision to support computational analysis of carving techniques, fracture surfaces, and fragment joins. Integrating these digital

representations with fragment metadata and historical documentation will allow the project to test how CIDOC-based models can represent relationships among physical objects, digital surrogates, and analytical observations.

This theme will also address the integration of historical catalog records and identifier systems associated with the fragments. Because the Marble Map fragments have been cataloged repeatedly across different scholarly projects, the corpus includes multiple numbering systems and historical references. The project will explore strategies for linking these identifiers within the knowledge structure so that fragments can be traced across successive cataloging efforts and scholarly interpretations.

Outputs of this theme will include:

- prototype knowledge structures representing portions of the Marble Map dataset
- integrated records linking fragments to digital scans, catalog entries, and historical visual representations
- documentation of data modeling approaches and implementation challenges

These prototypes will provide the experimental environment necessary for evaluating the project's research questions.

Research Theme 3: Analytical and Computational Research Workflows
Primary focus: Months 10–18

The third research theme investigates how analytical and computational research data can be incorporated into CIDOC-based knowledge structures. This work expands the experimental scope of the project by integrating scientific analysis and computational research outputs with the descriptive and historical data explored in earlier themes. This theme addresses the project's third and fourth research questions by exploring how analytical datasets and computational methods—including AI-assisted analysis of scholarly literature—can be incorporated into CIDOC-based knowledge environments.

One component of this theme focuses on incorporating scientific analyses associated with the fragments, including chemical analysis, spectrometry, and imaging data. These datasets often describe measurements associated with particular fragments, surfaces, or architectural features and therefore require modeling strategies capable of representing relationships between empirical observations and physical objects.

A second component of this theme investigates methods for transforming natural-language scholarship into structured data representations. Using AI-assisted approaches to text analysis and image recognition, the project will explore how references to fragments, images, and catalog identifiers appearing in scholarly publications can be automatically identified and linked to corresponding entities within the knowledge structure.

These experiments will allow the research team to evaluate how cultural heritage data infrastructures can incorporate both analytical research data and narrative scholarly documentation within a shared semantic framework.

Outputs of this theme will include:

- experimental integration of analytical datasets within the prototype knowledge structures
- evaluation of AI-assisted workflows for identifying object references within scholarly publications
- documentation of modeling approaches for linking computational and textual research outputs to structured cultural heritage data

Research Theme 4: Scholarly Consultation, Evaluation, and Dissemination
Primary focus: Months 16–24

The final research theme focuses on evaluating the modeling approaches developed during earlier themes and disseminating the project's findings to the broader library and cultural heritage communities. This theme provides the evaluation framework through which the project's modeling approaches will be tested and refined through collaboration with external scholars and institutions.

Scholarly consultation is a central component of this theme. At the end of the first year of the project, the team will host a directed working meeting at Indiana University Indianapolis that will bring together invited collaborators from partner institutions. This workshop will focus on examining the developing ontological framework and evaluating how it aligns with related digital humanities and cultural heritage data initiatives.

Later in the project period, the team will host a symposium that brings together scholars, librarians, technologists, and cultural heritage professionals to discuss the outcomes of the research and the broader implications for integrating emerging cultural heritage data types into library and archival infrastructures. These events will function not only as dissemination activities but also as opportunities to evaluate the applicability of the project's methods across

different research domains.

The project team will also prepare publications, documentation, and modeling guidelines describing how CIDOC-CRM–based approaches can support research-oriented cultural heritage data environments. These materials will be made publicly available to support adoption and further experimentation by libraries and cultural heritage institutions.

Outputs of this theme will include:

- publicly available modeling guidelines and documentation
- scholarly publications and conference presentations
- workshops and symposia engaging cultural heritage and library communities
- recommendations for applying CIDOC-based approaches to complex research datasets

Expected Outputs of the Project

Across these four research themes, the project will produce:

- conceptual frameworks for representing research-oriented cultural heritage data
- prototype knowledge structures integrating heterogeneous datasets
- experimental workflows linking analytical data, computational research, and scholarly literature
- documentation and modeling guidelines for libraries managing complex digital research datasets
- scholarly publications, workshops, and symposia that disseminate the project's findings

Together, these outputs will contribute new methodological approaches for representing complex cultural heritage research environments within library information infrastructures and will advance the use of CIDOC-CRM as a framework for research-oriented cultural heritage data modeling.

PROJECT RESULTS

Advancing Cultural Heritage Research Infrastructure

This project will produce new conceptual and practical approaches for representing complex cultural heritage research data within library information infrastructures. By investigating how CIDOC-CRM can support knowledge structures that integrate digital representations, scientific analysis, computational research, and scholarly interpretation, the project addresses a growing challenge facing libraries and archives that steward digital research datasets.

The primary result of the project will be a set of documented approaches for representing research-oriented cultural heritage data using CIDOC-CRM–based modeling strategies. These approaches will be tested through prototype implementations using the Marble Map dataset, which includes high-resolution three-dimensional scans of archaeological fragments, analytical measurements, historical documentation, and scholarly literature. Through these experiments, the project will evaluate how heterogeneous datasets associated with cultural heritage objects can be integrated within shared semantic structures.

The findings will contribute new guidance for libraries and archives seeking to organize and connect complex research datasets. As digital humanities and computational archaeology increasingly rely on analytical data and digital models, the ability to represent these materials within library systems will become increasingly important. By exploring the capabilities and limitations of CIDOC-CRM in this context, the project will help inform future cultural heritage data infrastructures designed to support advanced research.

Modeling the History of Cultural Heritage Knowledge

A second major outcome of the project will be the development of approaches for representing the historical evolution of scholarly knowledge about cultural heritage objects. The Marble Map corpus illustrates how archaeological datasets often exist not only as collections of physical artifacts but also as layered records of scholarly interpretation. Over the past three centuries the fragments of the Marble Map have been cataloged, drawn, engraved, photographed, and digitally modeled in a variety of forms. Each of these representations reflects different scholarly assumptions and reconstruction hypotheses.

The project will therefore explore how CIDOC-based knowledge structures can represent the historical processes through which archaeological corpora are documented and interpreted. By linking fragments to their historical textual, visual, and digital representations, the project will demonstrate how libraries can make the history of scholarship itself accessible as structured data.

This approach will allow researchers to trace how interpretations of the Marble Map have changed over time, compare competing reconstructions, and examine how fragments have been represented across different publications

and cataloging systems. Modeling these relationships will provide new ways of studying the development of archaeological knowledge and will demonstrate how library information systems can support the documentation of scholarly processes.

Integrating Analytical and Computational Research Data

The project will also generate insights into how scientific and computational research data can be incorporated into cultural heritage knowledge structures. Analytical techniques such as chemical analysis, spectrometry, and imaging increasingly play central roles in the study of archaeological materials. At the same time, computational approaches—including machine learning methods used to analyze fragment geometry and potential joins—are producing new forms of research output.

These analytical and computational results often remain disconnected from library metadata systems because they are generated through research workflows rather than traditional cataloging processes. Through prototype implementations developed during the project, the research team will evaluate how these datasets can be represented and connected within CIDOC-based models.

In addition, the project will investigate experimental workflows for integrating information extracted from the scholarly literature associated with the Marble Map corpus. Using AI-assisted techniques, the project will explore methods for identifying references to fragments, images, and catalog identifiers appearing in archaeological publications and incorporating those references into the knowledge structure. These experiments will provide insights into how libraries might integrate textual scholarship with structured research data in the future.

National Impact for Libraries and Cultural Heritage Institutions

The outcomes of this project will benefit libraries, archives, and cultural heritage institutions that manage complex research datasets. Many institutions are now responsible for stewarding digital models, analytical measurements, and computational research outputs alongside traditional archival materials. However, widely adopted metadata frameworks have not yet fully addressed the challenges of representing these forms of data.

By evaluating how CIDOC-CRM can support research-oriented cultural heritage knowledge structures, the project will provide libraries with new conceptual tools for organizing and connecting heterogeneous research datasets. The modeling guidelines, prototype implementations, and documentation produced by the project will be made publicly available so that other institutions can build upon the results.

These resources will help libraries support emerging forms of digital scholarship that rely on integrated datasets and computational analysis. The project will therefore contribute to the development of information infrastructures capable of supporting the next generation of cultural heritage research.

Dissemination

The results of the project will be shared broadly with library, cultural heritage, and digital humanities communities. The research team will publish project findings through scholarly articles, conference presentations, and publicly available documentation describing the modeling approaches developed during the project.

The team will also engage with professional communities through workshops and presentations at conferences related to digital humanities, library and information science, and cultural heritage data management. These activities will provide opportunities for librarians, archivists, and researchers to evaluate the project's findings and to explore how similar approaches might be applied to other cultural heritage datasets.

Through these dissemination efforts, the project will encourage further experimentation with research-oriented cultural heritage knowledge structures and will contribute to ongoing conversations about how libraries can support complex digital research environments.

Schedule of Completion

The schedule of completion translates the project's four research themes into a coordinated sequence of activities designed to address the research questions and produce the project's intended results. Year 1 emphasizes Theme 1 (Conceptual Modeling) and Theme 2 (Data Integration), during which the team will analyze the Marble Map corpus, develop CIDOC-CRM conceptual models, and implement prototype knowledge structures linking fragments, digital representations, and historical catalog systems. These activities address Research Questions 1–2 and generate the project's initial modeling framework and prototype datasets. Year 2 focuses on Theme 3 (Analytical and Computational Workflows) and Theme 4 (Scholarly Consultation, Evaluation, and Dissemination). During this phase, the team will integrate analytical and computational research data, test AI-assisted extraction of references from scholarly publications, and evaluate the resulting knowledge structures through collaboration with external scholars and cultural heritage partners. These activities address Research Questions 3–4 and culminate in the project's primary results: interoperable prototype knowledge structures, publicly available modeling guidelines, scholarly publications, and workshops designed to disseminate findings and support adoption by libraries and cultural heritage institutions.

Year 1: September 2026 – August 2027

Period	**Project Activities and Benchmarks**
July–August 2026 (pre-award)	Partner coordination meetings to confirm data access, technical infrastructure, and project management protocols supporting **Themes 1–2**.
September–October 2026	Project initiation (**Theme 1**). Final confirmation of research questions and conceptual modeling goals. Project initiation (**Research Question 1**).
September–November 2026	Comprehensive analysis of data categories associated with the Marble Map corpus (**Theme 1**).
October–December 2026	Development of initial CIDOC-CRM conceptual models representing relationships among data categories (**Theme 1**, addressing **Research Questions 1–2**).
November 2026 – February 2027	Evaluation of relevant CIDOC-CRM extensions (CRMsci, CRMarchaeo, CRMdig) to determine modeling strategies (**Theme 1**, **Research Question 3**).
January – April 2027	Development of prototype schema structures integrating fragment metadata, digital scans, and historical identifier systems within a CIDOC-based structure (**Theme 2**).
February – June 2027	Initial implementation of prototype knowledge structures linking fragments to digital representations and historical catalog records (**Theme 2**, addressing **Research Questions 1–2**).
March – August 2027	Integration of high-resolution 3D scan data and associated metadata into the prototype CIDOC-based data model (**Theme 2**, **Research Question 3**).
May – August 2027	Documentation of conceptual modeling strategies and preparation of internal technical reports describing modeling approaches and implementation challenges (**Theme 2**).

Year 1 Benchmark

- Completion of CIDOC-CRM conceptual modeling framework (**Theme 1 output**)
- Prototype knowledge structures integrating fragment metadata, digital models, and historical catalog systems (**Theme 2 output**)

Year 2: September 2027 – August 2028

Period	Project Activities and Benchmarks
September – November 2027	Expansion of prototype knowledge structures to incorporate analytical datasets associated with Marble Map fragments, including imaging data and scientific measurements (**Theme 3**, **Research Question 3**).
October 2027 – February 2028	Development of modeling strategies for representing computational research outputs derived from 3D models, including geometric descriptors and fragment comparison metrics (**Theme 3**).
November 2027 – March 2028	Implementation and testing of AI-assisted workflows for identifying references to fragments, images, and catalog identifiers in scholarly publications (**Theme 3**, addressing **Research Question 4**).
January – April 2028	Integration of extracted scholarly references into CIDOC-based knowledge structures linking fragments, digital representations, and scholarly publications (**Theme 3**, **Research Question 4**).
February – May 2028	Evaluation and refinement of modular knowledge structure design based on experimental results and collaboration with project partners (**Theme 4**).
March – June 2028	Preparation of modeling guidelines describing how CIDOC-CRM can support research-oriented cultural heritage data infrastructures (**Theme 4 output**).
April – June 2028	Organization of scholarly workshop at Indiana University Indianapolis bringing together collaborators and specialists in cultural heritage data modeling (**Theme 4**).
June 2028	Directed working meeting evaluating the project's ontological framework and its interoperability with related digital humanities and cultural heritage initiatives (**Theme 4**).
June – August 2028	Preparation of scholarly publications, technical documentation, and public-facing project reports describing modeling strategies and research results (**Theme 4 output**).
July – August 2028	Final project evaluation and preparation of final reports for IMLS (**Theme 4**).

Final Project Benchmark

- Completed prototype knowledge structures integrating artifacts, digital models, analytical datasets, and scholarly references
- Publicly available modeling guidelines and documentation
- Scholarly workshop and publications disseminating project results