White Paper

Report ID: 111933
Application Number: HD-51921-14
Project Director: Michael David Carrasco
Institution: Florida State University
Reporting Period: 5/1/2014-9/30/2015
Report Due: 12/31/2015
Date Submitted: 3/7/2016
White Paper Report

The Mesoamerican Corpus of Formative Period Art and Writing

Grant # HD-51921-14

Project Directors:

Michael D. Carrasco
Department of Art History, Florida State University

Joshua D. Englehardt
Centro de Estudios Arqueológicos, El Colegio de Michoacán

Dennis E. Slice
Department of Scientific Computing, Florida State University

Reporting period: 05/01/2014–09/30/2015

February 10, 2016
## Table of Contents

Abstract  4

Project Participants  5

Narrative Description  6

Project Rationale  7

Background: Olmec Iconography and Writing  7

Project Contribution  9

Environmental Scan  10

History and Duration of the Project  11

Activities and Achievements Realized during Grant Tenure  15

*Fieldwork and Preparation of Data and Digital Assets*  15

  Conservation  17

*Seriation, Sorting, and Classification*  17

  Iconographic analysis and the development of writing  18

*Digital Architecture and Applications*  19

  Technical Specifications  20

    Server Side  20

    Web Server (NodeJS)  21

    Client Side  21

  Shape Matching  22

  Design Evaluation  23

*Dissemination and Publications*  24

  Products derived from project activities  24

    Publications  24

    Professional presentations  25
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic courses derived from project work</td>
<td>27</td>
</tr>
<tr>
<td>Advising</td>
<td>28</td>
</tr>
<tr>
<td>Collaborations</td>
<td>28</td>
</tr>
<tr>
<td>Final Remarks</td>
<td>29</td>
</tr>
<tr>
<td>References</td>
<td>31</td>
</tr>
<tr>
<td>Appendices</td>
<td>37</td>
</tr>
</tbody>
</table>
Abstract

The purpose of *The Mesoamerican Corpus of Formative Period Art and Writing* is to create a digital resource on Formative period art and writing that facilitates research on the emergence of writing in the New World. To achieve this goal, during the tenure of the grant period the Corpus has assembled a collection of Formative period (ca. 1500–400 BCE) Mesoamerican iconography and writing, constructed a prototype database architecture, and designed an intuitive, interactive interface targeted to a broad range of users. Start-Up phase funding allowed us to build the project’s core database, generate digital assets, and plan for the incorporation of visual recognition software into mobile and web applications to be developed during the next phase of the project. Upon its completion this project will construct a critical resource for scholars of the Formative period Americas and extend and develop technologies and data management architectures that will have wide application in art history, archaeology, and kindred humanities disciplines. Our long term goal is the creation of a sophisticated, bilingual digital asset and software toolkit that presents the full richness of Formative period art and writing to a broad audience and introduces new ways of conceptualizing and approaching visual data in humanities research.
Project Participants

**Co–Director:** Michael D. Carrasco, Associate Professor of Art History, Florida State University

**Co–Director:** Joshua D. Englehardt, Research Professor of Archaeology, Centro de Estudios Arqueológicos, El Colegio de Michoacán

**Co–Director:** Dennis Slice, Associate Professor of Scientific Computing, Florida State University

Cameron Berkley, Database and Software Developer, Graduate Research Assistant, Scientific Computing Morphometrics Lab, Florida State University

Roberto Lunagómez Reyes, Curator of Olmec collections, Museo de Antropología Xalapa, Universidad Veracruzana

Rebecca Perales Vela, Director, Museo Regional de Antropología Carlos Pellicer Cámara, Instituto Estatal de Cultura del estado de Tabasco
Narrative Description

This project created the digital architecture, database, and computational tools to support *The Mesoamerican Corpus of Formative Period Art and Writing*, a digital resource that will expand currently available data to facilitate and robustly support research on the emergence of writing in the New World by developing a suite of digital research tools. The goal of this on-going project is to reveal the artistic and scribal traditions of the Olmec culture, Mesoamerica’s first civilization and the originator of an ancestral sign system from which all later Mesoamerican writing developed. To work to this goal, the *Corpus* assembled a database of Formative period (ca. 1500–400 BCE) Mesoamerican iconography and writing in a digital catalogue accessible via intuitive, interactive applications and targeted to a broad range of users, from academic researchers to K–12 teachers and students. Start–Up funding has allowed us to build the project’s core database, assemble digital assets, and plan for the incorporation of visual recognition software into mobile and web applications to be developed during the implementation phase. Digital tools based on emerging technologies will allow for innovative ways of searching—such as visual input searches—comparing, analyzing, and visualizing relationships among a broad corpus of visual data, archaeological materials, and contextual information. This project both constructs a critical resource for scholars of the Formative period Americas and extends and develops technologies and data management architectures that will have wide application in art history, archaeology, and kindred humanities disciplines. Our long-term goal is the creation of a sophisticated, bilingual digital asset and software toolkit that presents the full richness of Formative period art and writing to a broad audience and introduces new ways of conceptualizing and approaching visual data in humanities research. Our work will facilitate the exploration of Formative period Mesoamerican art and writing, develop new technologies for visual searches valuable to the study of visual culture more generally, and achievements to date have built a solid foundation for the implementation of the long–range goals of the *Corpus* project.
**Project Rationale**

Investigations of the origins of writing and decipherments of unknown scripts rely on a comprehensive corpus of material subjected to analysis from a broad range of fields. This project creates just such a catalogue for Formative period Mesoamerican iconography and writing, which has been understudied. The Olmec corpus is unusually rich, but widely scattered, and access to data is frequently restricted. Further, some scholars have questioned the place of the Olmec sign system in the history of writing in the New World, arguing that the Olmec iconographic system was of insufficient visual and narrative complexity to provide the basis for Mesoamerican writing (Houston 2004). However, as Schmandt–Besserat (2007, 2010) has shown, art played a vital role in the emergence of the cuneiform script in the Near East. This project traces a similar trajectory in Mesoamerica, in which Formative period Olmec iconography played a crucial role in the development of all Mesoamerican scripts.

Exciting new discoveries and four decades of intensive research have provided a clear view of the steps spanning art and writing in Mesoamerica that potentially illustrate the critical moment when writing was invented. This project will significantly advance research on this period of transition from iconography to phonetic scripts and thereby test current hypotheses on the emergence of writing in one of the few locations in the world where writing developed. By combining new archaeological finds with poorly documented data gleaned from disparate, restricted sources into a centralized database, and creating a suite of analytic digital tools accessible through web and mobile device applications, our project will allow investigators to consider the evidence in a new light, to generate new hypotheses, and to reassess the critical role of Olmec art in the invention of writing in the Americas.

**Background: Olmec Iconography and Writing**

The Olmec are known both as a Formative period culture centered along the coast of the Gulf of
Mexico and as an iconographic assemblage and artistic style that stretched from western Mexico to Costa Rica from approximately 1500–400 BCE. The recent excavations of the Cascajal Block (Rodríguez Martínez et al. 2006) and the San Andrés roller stamp (Pohl et al. 2002), dating to 900–600 BCE, isolate the emergence of Mesoamerican writing in the Olmec heartland. The antiquity of Olmec culture and the widespread presence of Olmec iconographic elements in subsequent traditions suggest that later regional writing systems, such as Mayan hieroglyphs, developed from a Formative period Olmec iconographic and sign system. Accordingly an understanding of this foundational system is vital for comprehending many of the hallmark features of later Mesoamerican cultures, in which writing played a significant social role as a medium of artistic expression and a communicative technology.

Current scholarship on script emergence in Mesoamerica (Houston 2004; Justeson 1986; Pohl et al. 2008) suggests that the critical transition from iconography to writing involves the “divorce” of iconic elements from a purely visual frame of reference and the subsequent recontextualization of signs or motifs within a grammatical, linguistic framework. As iconography moves towards writing, the organization and interpretation of signs conforms to the structure of language, as opposed to the representational schemes of which a given sign was originally an element. To understand how and why signs were pulled from iconographic contexts and incorporated into the structure of writing, researchers must focus on the interplay between the different contexts in which specific signs were deployed.

To unravel these contexts, investigators need a complete catalogue and searchable database of Formative period iconographic and archaeological data that goes beyond the sole existing catalogue (Joralemon 1971), which lacks the wealth of new evidence that has come to light since its publication. Selective studies of Olmec iconography (e.g., Clark and Pye 2000; Coe 1965; Quirarte 2007; Reilly 1991, 1996; Taube 2004) are similarly incomplete and are limited to specific regions or periods. Therefore despite decades of investigation, our understanding of the origins and development of Mesoamerican writing remains limited. To
address this problem our project weds this past material with unpublished objects and recent discoveries to build a comprehensive database. Through the use of innovative digital tools, the primary goal of this project is precisely to track the development of specific signs during the Formative period so as to visualize the recontextualizations of iconographic motifs in ways that will most advance research on Formative period art and script development.

**Project Contribution**

Our proposal takes a holistic view, acknowledging that both scholarly and technological progress is needed to elucidate Formative period iconographic systems and script emergence in the ancient New World. This project addresses these lacunae in four ways:

1. By providing a comprehensive corpus that unites limited existing catalogues with the large collection of unpublished or partially published Formative period iconography and writing in one easily searchable location through an intuitive and interactive digital resource;

2. By producing an initial chronological seriation and spatial contextualization of Formative period signs tied to firmly dated, scientifically excavated material that allow researchers to visualize and generate hypotheses about the spatial and temporal distribution of Formative period iconography;

3. By outlining methodologies for approaching the origins and development of Mesoamerican scripts that have generally been overlooked by researchers; and

4. By extending and developing digital assets and computational tools that facilitate the visualization of complex relationships otherwise difficult to express among diverse datasets.
During this grant tenure, we have approached these goals by surveying new technologies based on advances in visual recognition software, including shape context description (Belongie et al. 2002; Hyder et al. 2009; Roman–Rangel et al. 2009), optical character recognition and pattern analysis (Bataineh et al. 2012; Batuwita et al. 2011; Dalitz et al. 2008; Smith 2007), and polymorphic shape rule–based detection (Frauel et al. 2006), as well as probabilistic machine learning and computer vision, to extend and develop original digital tools for visual input searches and the presentation and comparison of data. Given the inherent visual nature of the material, searches that are based on the input of images—such as a researcher’s own drawing of a motif—will permit investigators to explore the corpus of Formative period Mesoamerican art and writing in ways impossible with keyword searches. Just as whole document searches revolutionized the possibilities for text–based research and made available sources that were previously restricted to specialists, visual input searches will allow investigators to pose fundamentally different questions of visual datasets, and to create new ways of expressing their findings. The adaptation of the technologies used in this project also poses exciting possibilities in the humanities for those areas in which there is a need for the visualization of complex relationships among multiple datasets that are difficult to express in narrative alone.

Our applications and digital assets integrate computational tools capable of describing, identifying, and comparing complex pictorial elements and delivering varied iconographic, contextual, and archaeological data to many types of users. The comprehensive catalogue and searchable database of Formative period iconography will benefit research by greatly expanding resources currently available for research on the development of Mesoamerican writing, thereby making key contributions to the larger fields of art history, archaeology, Mesoamerican studies, and the history of writing.
Environmental Scan

Two NEH funded projects are similar to the *Corpus*. The University of Houston’s Visual Web Interface for Researchers (Vwire) project has produced an online tool that offers an intuitive way of comparing, categorizing, and sorting visual data. Vwire highlights the power of making visual arguments that are difficult or impossible to achieve through narrative. The second, FACES: Faces, Art, and Computerized Evaluation Systems, based at the University of California, Riverside, utilizes facial recognition software to identify artistic portraits that lack provenience. Our project will incorporate similar technologies to help isolate patterns in Formative period art and writing, as well as present this material in ways that, like Vwire, enable the visualization of relationships between multiple datasets.

In stage three of this project, we plan to adapt the pattern recognition technologies and methodologies for visual searches employed in projects such as Detexify (http://detexify.kirelabs.org) and *Leafsnap: An Electronic Field Guide* (http://leafsnap.com). Detexify was developed to help those working with the document markup language LaTeX to find specific symbols. It allows users to input visual information, in this case as a drawing, from which a list of similar signs is displayed. Likewise, *Leafsnap* is designed to recognize plant species from photographs taken by the user by comparing images against the Smithsonian Institution’s large centralized specimen collection. Such a feature allows users with no specialized botanical knowledge to access detailed information on taxonomy and species identification. User data are automatically incorporated in the database, thereby expanding the dataset as the program is used. For our project, the Museo de Antropología de Xalapa, Veracruz and the Museo Regional de Antropología Carlos Pellicer Cámara in Villahermosa, Tabasco, which house the largest collections of Olmec art in the world, serve a similar function to make our corpus the largest digitally accessible collection of Formative period iconography and writing.
History and Duration of the Project

Our research is a long term, multi–year project that has grown out of a fusion between Carrasco’s research on Mesoamerican art and epigraphy and Englehardt’s focus on interregional interaction in the Formative Period and the evolution of script traditions in the Americas (Carrasco 2012a, 2012b; Carrasco and Englehardt 2011; Englehardt 2010, 2012). This project unfolds in three stages, the first of which was a 2012–2013 phase of planning and documentation, funded by Florida State University. During this initial period of support, the co–directors traveled to Mexico in summer 2012 to review collections at the Museo de Antropología de Xalapa and the Museo Carlos Pellicer Cámara. We secured agreements with these institutions to ensure access to their restricted collections. We documented a limited range of material at these museums and regional archaeological sites through high–resolution digital photography and constructed a preliminary database of approximately 150 artifacts that provides chronological, contextual, archaeological, and descriptive information on each object. We also investigated the best practices for the preparation of imagery so that it can be subjected to pattern recognition technologies that will permit both visual searches of the database and multi–channel methods of dissemination. By the end of stage one, we had identified a range of appropriate technologies for content management and online delivery systems.

During stage two, from May 2014 through September 2015, we constructed and populated a core database and continued the analysis of the most important motifs in our preliminary corpus. Carrasco and Englehardt conducted a further four months of museum research in Mexico over 2014 and 2015. Additional fieldwork added a significant number of poorly documented Formative period objects to the database. We generated a preliminary stylistic seriation and cultural–geographical contextualization of the corpus to refine the current spatial and chronological frameworks for Formative period art and writing.

Funding also supported collaboration with Dennis Slice of the Department of Scientific
Computing at FSU, who oversaw the research and preliminary development of the technological infrastructure necessary to display the spatial and temporal relationships among signs and motifs within the database, and to represent the complex regional synchronic and diachronic relationships among datasets. Slice’s Morphometrics Lab (http://morphlab.sc.fsu.edu/) is a leading center for the development of analytical methods and software for the characterization, comparison, modeling, and manipulation of such complex, multi–dimensional data, thereby providing a solid understanding of and proven experience in the application of existing tools on conceptual and practical levels. During the grant period, Slice and software developer Cameron Berkley evaluated available resources and visual search routines, specified data structures according to field–specific best practices, created preliminary wireframes of web–based and mobile interfaces, and developed prototype database digital architecture.

The evaluation of open–source pattern recognition technologies will permit visual searches to be incorporated during stage three of the project. Therefore, a crucial component of the second stage was Carrasco and Englehardt’s preparation of the data and digital assets to take advantage of these technologies. These include iconographic data documents that are comprised of layers of high–resolution, digital photographs collected through fieldwork, historical photographs, and line drawings created by the project. Such drawings will segment individual motifs from a composition onto their own layer to facilitate high–speed retrieval of specific iconographic elements. Separating multiple icons that co–occur in a single composition will allow for the comparative search for relationships among signs. This is a critical search ability for understanding an iconographic element’s distribution in art and incipient writing systems. These composite images will eventually permit visual input searches—both of entire compositions and single elements within compositions—against images in the database.

Finally, using the technologies described above, and integrating GIS–based interactive maps and excavation data, we developed a database that serves as the repository for the digital resources that feed the mobile and web applications to be developed in stage three of the
project. We generated proof of concept prototypes and wireframes of the application interfaces, including: 1) an initial website providing broad access to project data; and 2) wireframes of web-based and mobile applications (e.g. iOS or Android) that will eventually provide access to core digital assets through a high quality, intuitive user interface. In sum, we have constructed a preliminary digital database of Formative period iconography; evaluated, developed, and tested backend software architecture; produced digital assets and prepared data to make full use of new technologies; and created initial designs for the database’s user interface. The beta database, initial digital architecture and software development, and creation of digital assets will form the basis for a proposal for an NEH Digital Implementation Grant. Future funding for stage three will allow for the creation of a mobile application and web-based desktop client, the implementation of visual input searches, and a major expansion of the catalogue of objects and contextualizing information. During stage three of the project, we will incorporate expanded functionalities such as visual input searches based on pattern recognition software into fully operational web-based and mobile applications. In addition, we will explore methods that allow users to input their own material into the database, thus further expanding the corpus while preserving the academic integrity of the product.
Activities and Achievements Realized during Grant Tenure

During the grant period we have successfully realized our stated goals within the timeframe outlined in the original work plan, incorporated new documentary techniques, such as Reflectance Transformation Imagery (RTI)\(^1\) and 3D photogrammetry,\(^2\) and disseminated our findings and methods through publications, presentations, and courses. Additionally, we have isolated database technologies and structures that will enable the project to build in many of the search capabilities outlined in the proposal, and add viewing options, such as 3D objects, which we had not initially anticipated. The database structure will also support the eventual implementation of image input or reverse image search capabilities that we hope to include in future iterations of the project as described in the proposal. In sum, project work has produced a core set of digital assets and facilitated the investigation of computational architecture that will allow for the eventual development of the fully functional website and application envisioned in our original proposal. These foundational steps will serve as the basis for our Digital Humanities Implementation Grant proposal, currently in development.

Fieldwork and Preparation of Data and Digital Assets

During the tenure of the grant co-directors Carrasco and Englehardt conducted two major periods of fieldwork (August-September 2014 and March 2015) in conjunction with project advisors Roberto Lunagómez Reyes and Rebeca Perales Vela. The first season of fieldwork focused on the collections at the Museo de Antropología de Xalapa (MAX), Veracruz, Mexico. There, we documented approximately 15 previously unrecorded Formative period archaeological objects with iconography. We created high-resolution digital images of these and approximately 75 other objects in MAX collections and also applied Reflectance Transformation Imagery (RTI)\(^1\) and 3D photogrammetry,\(^2\).


Imaging (RTI) and 3D Photogrammetry documentary technologies on these objects (see previous interim reports; see Appendices 1–3). MAX also generously provided 164 of their own high resolution images of objects in their collections and allowed us to incorporate these images into the project’s multimedia and digital databases.

During the second stage of fieldwork in March 2015 we documented objects at the Museo Regional de Antropología Carlos Pellicer Cámara and the Parque-Museo La Venta in Villahermosa, Tabasco, Mexico. There, we documented approximately 50 Formative period archaeological objects with iconography, including approximately 10 previously unrecorded objects. We created high-resolution digital images of these, as well as RTI and 3D images. Our imaging techniques revealed new details on many of these objects, some of which had not been identified or published previously or had been lost due to erosion or deterioration of the monuments (see Appendices 4-10). The incorporation of details difficult to discern either through naked-eye inspection or in conventional photographs and absent in most previous drawings were one of the initial impulses for the project. The discovery of such details confirms the value and utility of documenting sculpture and inscriptions through a variety of imaging technologies. These details greatly enhance the final project drawings.

Fieldwork over the course of the grant tenure allowed for the production of high-quality digital assets, including RTIs, 3D models, vector line drawings, and multimedia materials, as outlined in the original work plan (tasks 2–3). In addition, RTI and 3D photogrammetry, as cost-effective methods, allowed us to produce results comparable to those produced by, for example, commercial 3D scanners, at a fraction of the cost. As detailed below, we have also been able to explore possibilities for including open-source visualization applications for RTI and 3D photogrammetric renders in the final versions of the website and mobile application.

---

3 E.g., Meshlab (http://meshlab.sourceforge.net/)
Conservation

The discovery of details provoked a reexamination of historic photographs taken at the time of the artifact’s original excavation. Our reexamination revealed that a number of features are indeed present in historic photographs that have been “forgotten” and are now only clearly evident through RTI, raking-light, photogrammetry, or other techniques. The differences between historic photographs and the current condition of the objects suggests in some cases extensive erosion. Our “rediscovery” of details not often discussed in recent literature suggests that the techniques we have applied through this project may be profitably applied in other contexts in which artifact deterioration is a concern, thereby recovering or preserving vital information encoded on the objects and aiding in conservation efforts.

Additionally we hope to obtain copies of historic photographs of La Venta monuments from the Smithsonian Institution. From our work with the monuments, specifically those from the Parque-Museo La Venta decades of weathering have caused significant erosion. In the case of La Venta Monument 13, illustrated in Appendix 10, details found in the historic photograph now no longer exist on the actual object. For us this has reinforced the importance of compiling as much of the material on a given object as possible since without this history of documentation available it is difficult for researchers to judge preservation and the institutions responsible for these irreplaceable objects to monitor their condition over the long term.

Seriation, Sorting, and Classification

The drawings, multimedia materials, high resolution photographs, and other digital assets produced through fieldwork have been incorporated into the prototype database. Over the course of the grant tenure, we developed and refined informational fields, populating the database with a variety of descriptive and classificatory information on each object documented through fieldwork (e.g., name, dimensions, object type and class, contextual data, etc.). The

4 For some examples of their holdings see: http://anthropology.si.edu/olmec/english/index.htm
cataloging and incorporation of information from disparate sources into a centralized, unified
database was one of the primary goals of the project, and we have made significant advances
in terms of the description, seriation, and classification of over 150 Formative period objects. In
future project phases, we will expand this preliminary database, adding information as it
becomes available. The open source programs we employ (as detailed below) eventually will
also allow for individual users to upload information and multimedia files, continually updating
and expanding the corpus through the integration of new data.

Iconographic analysis and the development of writing

The seriation and classification of a wide variety of new and previously gathered data has also
greatly aided in the ongoing analysis of Formative period art and iconographic motifs that was a
major motivation for this project. As detailed below, our work has led to a number of new,
potentially valuable interpretations of Formative period art, which we have disseminated in
through variety of avenues. One area in which project work has allowed us to make significant
advances is in the investigation of the artistic conventions of Formative period art and incipient
scripts. Of particular note is our identification of previously understudied linguistic tropes such as
kennings in a variety of contemporaneous artistic objects. Such analysis would not have been
possible without the creation of a centralized corpus that is at the heart of this project. We are
confident that future project work will allow us and other researchers to make further important
advances in the field. Employed systematically, analyses like those that derive from project work,
that track the changing contexts and functions of signs against a broad corpus of visual culture,
will likely be one of the major means by which we can further the study of early examples of
writing for which we have but a single or few examples, thus contributing to a better theoretical
understanding of the origins and role of writing in early civilizations.
Digital Architecture and Applications

During the grant period we have evaluated digital database architecture, refined database fields, and developed a naming-numbering scheme for specific objects within the database and associated digital assets. In particular we have focused on developing the technical groundwork for the database and database driven services and UML mapping the relationships between informational items, database fields, and system components (see Appendices 11–12). A MongoDB\(^5\) database has been populated with a subset of corpus data in order to test the web and mobile prototypes. Both the website and the mobile application (when the mobile app is completed) are served data via a REST (Representational state transfer) framework implemented in NodeJS.\(^6\) This allows both services to access the database in the same underlying manner, while having different implementations by necessity. NodeJS is also responsible for serving the website itself to browsers, and several different modules work to route users to the correct page, handle file uploads, and automatically generate pages with different data on demand. We have identified OpenLayers as likely the best program for displaying map data with information queried from our corpus of data. All software packages used are free and open source.

We have designed and developed a database architecture and applications that link data to specific modes of visualization (i.e. the linking of the MongoDB database to a map interface using OpenLayers\(^7\); see Appendices 13–14). Below we provide brief descriptions of the technical specifications. The rich schema for data organization allows for the recording of the metadata associated with a range of data types from the objects that are the focus of this study to the drawings, photographs (both current and historical), Reflectance Transformation Images, and 3D models, among others, that supplement these prime objects. We have examined (and downloaded, where possible) the open-source codes and APIs (Application

---


Programming Interface) that eventually will permit for the seamless visualization of digital assets such as RTIs and 3D photogrammetric renderings via the website and mobile application. A subset of iconographic data documents and digital assets have been incorporated into the digital database architecture, and we will continue to upload assets as they become available. We continue to explore technologies and methods for reverse image searches, and are evaluating digital software and infrastructure so that the project digital assets may eventually utilize visual search functionality. As outlined in the original work plan, during the project we have thus conceptualized, refined, and created proof of concept digital wireframes that will serve as the foundation for our prototype web and mobile apps.

The digital architecture is based on open–source software and content management systems. Access to raw data, imagery, interpretive materials, and descriptive meta–data will be open and available for download in future iterations of the project. Project materials are licensed under a Creative Commons Attribution–Non Commercial 3.0 Unported License, permitting free access, use, and academic exchange of contents. The use, distribution, and modification of computational tools developed for the project will be made publicly available through the project website. The digital database is hosted through the FSU College of Fine Arts website.

Technical Specifications

The application can broadly be separated into two sections: the server side and the client side. The server side consists of a NoSQL MongoDB database and NodeJS web and data server. The client side consists of the user’s browser, running javascripts supplied by the web server.

Server Side

Database Server (MongoDB)

The corpus is stored using MongoDB a NoSQL database which contains “documents” in the form of Javascript Object Notation (JSON) key-value pairs. Several document types were
implemented, including Artifact, Sign, and Site (see Appendices 15–16). MongoDB has an extensive syntax for searching through documents.

**Web Server (NodeJS)**

NodeJS serves data to the client in the form of JSON via a REST (REpresentational State Transfer) API. This allows users to create, retrieve, update, or delete documents in the database by accessing URLs on the server (e.g., if the client requests the URL /db/artifacts/507f1f77bcf86cd799439012, the server will respond with a JSON object representing the artifact with the ID 507f1f77bcf86cd799439012). The Mongoose package\(^8\) is used to access MongoDB through NodeJS. The user does not generally access the REST API directly; rather it is accessed by the client (web browser) through a user interface.

NodeJS is also responsible for serving client-side web pages and associated javascripts. The Express package\(^9\) is used to route requests to URLs to the appropriate content. The Jade templating engine\(^10\) is used to allow for code re-use.

To prevent malicious or accidental deletion or modification of the corpus by end users, the Passport package\(^11\) is used to restrict certain actions on the REST API to authorized users only.

**Client Side**

The client side consists of html and javascript, provided by the server, which is displayed and run in the user’s web browser. The client is capable of asynchronously requesting data from the server REST API and displaying it. This allows an interactive experience without requiring the page to be refreshed each time the user wants to see new data.

---


The Bootstrap javascript framework\textsuperscript{12} was used to provide responsive design elements such as buttons and sliders. jQuery\textsuperscript{13} made the asynchronous calls to the REST API possible.

The OpenLayers3 interactive mapping library was extensively used to display the location and layout of artifacts and sites. Combining OL3 and jQuery, the user may interactively select artifacts and sites to see more information about them.

**Shape Matching**

Content-based image retrieval is the task of querying a database of images to find those images which are most visually-similar to a target image. The similarity of two images is assessed based on the contents of the images, including color, texture, and shape. A broad number of algorithms have been developed to accomplish this task (Marshall and Gunasekaran 2014). However, the lack of color and texture in Olmec sign tracings, and a small sample size, preclude the use of many of these algorithms. Researchers of Mayan signs have in recent years found success with a model known as Bag of Visual Words. Given the similarity of Mayan and Olmec signs, this model may be promising.

Bag of Visual Words is based on an earlier computational linguistics model known as Bag of Words. Bag of Words represents a document as the histogram over its vocabulary. The similarity of two documents can then be assessed as the distance between their bags of words (Salton and McGill 1986). Bag of Visual Words accomplishes the same task with images by defining a “visual vocabulary” of image feature descriptors. Image retrieval algorithms using Bag of Visual Words differ mainly in what feature descriptors they use, and how the basis of the “visual vocabulary” is defined.

One early feature descriptor which encodes shape information is the shape context. The shape context is a log-polar histogram over the distribution of points surrounding a given point.

\textsuperscript{12} Bootstrap · The world’s most popular mobile-first and responsive front-end framework (n.d.) Retrieved December, 2015 from \url{http://getbootstrap.com/}.

\textsuperscript{13} jQuery. (n.d.) Retrieved December 28, 2015 from \url{https://jquery.com/}. 
This captures shape information from the image while remaining invariant to rotation, translation, and scale. Roman-Rangel et al. (2011) proposed an extension of the shape context called the Histogram of Orientation Shape Context (HOOSC), which takes information about the gradient of the contours in an image. HOOSC proved to be more robust to the rough contours found in their dataset of Mayan glyphs. Rui and Collomosse (2013), working on sketch-based image retrieval, introduced the Gradient Field–Histogram of Oriented Gradients shape descriptor (GF-HOG). GF-HOG extrapolates a dense orientation field and computes the histogram of the gradient over this field. This helped overcome the sparse nature of sketches, and has found relevance in Mayan glyph retrieval (Roman-Rangel et al, 2013).

Once the visual feature descriptors are computed for each object in the database, a basis for the “visual vocabulary” must be established. Many researchers in the literature accomplish this by k-means clustering of the feature descriptors (Rui and Collomosse, 2013; Roman-Rangel et al. 2013). However more recent research by Chavan and Shahane (2015) used mean-shift clustering, which is less sensitive to initial conditions and is more robust to outliers.

With a basis for the “visual vocabulary” established, the bag of visual words for each sign in the database may be computed. When a user queries the database with an unknown sign, its bag of visual words is computed and compared with those in the database. The most similar sign, or most similar n signs, can then be identified and returned to the user.

While shape matching and visual input search capabilities remain to be developed in future iterations out initial exploration and evaluation of these technologies suggest that within the next several years developments in pattern recognition software will be sufficiently robust to enable us to apply these methods to iconographic issues. In Germany Dr. Nikolai Grube is heading a team of Maya epigraphers who are investigating similar technological issues. We hope that in the future synergistic collaboration with this and similar projects will bring reverse image searches into broader use in fields where visual research is a key component.
**Design Evaluation**

At various points throughout the design and functionality testing process, we have solicited feedback, both formally and informally, from a variety of colleagues, incorporating their suggestions (where possible and appropriate) into the final prototype designs.

**Dissemination and Publications**

Aspects of this project have been presented at a variety of venues from its brief, initial presentation at the grantee meeting in Washington, D.C. to professional meetings in the US and Mexico. We have also organized a session for the World Archaeological Congress in Kyoto, Japan, which focuses on the visualization of the archaeological object. We have published an article in the *Cambridge Archaeological Journal* and are preparing two additional publications. In addition to these formal presentations and publications of material we have circulated imagery generated from the project to a number of scholars focused on the Formative Period, soliciting feedback on the utility and design of project products. Finally, co-directors Carrasco and Englehardt are preparing a prospectus for a book on the origins and development of Mesoamerican scripts, which will be submitted to the University of Texas Press by the end of 2016.

**Products derived from project activities**

**Publications**

Carrasco, Michael D., and Joshua D. Englehardt

2015  

Englehardt, Joshua D., and Michael D. Carrasco

2016  
*El Corpus Mesoamericano del Arte y Escritura del Periodo Formativo: Informe*
Final del Proyecto Fase II. Report submitted to the Instituto Nacional de Antropología e Historia, Consejo de Arqueología, Mexico City.

Forthcoming
Englehardt, Joshua D., and Michael D. Carrasco

Carrasco, Michael D., and Joshua D. Englehardt

Berkley, Cameron, Dennis E. Slice, Michael D. Carrasco, and Joshua D. Englehardt

Berkley, Cameron, Dennis E. Slice, Michael D. Carrasco, and Joshua D. Englehardt
Professional presentations

Carrasco, Michael D. and Joshua D. Englehardt


Carrasco, Michael D.

2016 The Road to Cascajal: Digital Technologies in the Documentation of Olmec Art and Writing. Paper to be presented at Eastern Connecticut State University, February 25.

Carrasco, Michael D.


Carrasco, Michael D., and Joshua D. Englehardt


Englehardt, Joshua D., Michael D. Carrasco, and Mary D. Pohl


Englehardt, Joshua D., and Michael D. Carrasco


Forthcoming

Englehardt, Joshua D., Michael D. Carrasco, and Mary D. Pohl

2016  “New Perspectives on Gulf Coast Olmec Iconography and Scripts via the Mesoamerican Corpus of Formative Period Art and Writing." Paper to be presented at the 81st Annual Meeting of the Society for American Archaeology, Orlando, April 9.

Academic courses derived from project work

Carrasco, Michael D.

2015  Digital Documentation and the Preservation of Cultural Heritage. Summer course taught in conjunction with the Florida Department of State Collections.

Englehardt, Joshua D.

2016  Los Orígenes y el Desarrollo de la Escritura en Mesoamérica. Short course developed for el programa de movilidad universitaria, en la Secretaría Académica
de Historia de la Dirección General de CCH, UNAM.

2014  “Texto e Imagen en Sistemas de Escritura Mesoamericanos.” Invited lecture delivered to Centro de Estudios de las Tradiciones (El Colegio de Michoacán) graduate course Textos Amerindios I (Dr. Hans Roskamp). Zamora, October 31.

2014  “Introducción a la Escritura: Sus Orígenes y Desarrollo.” Invited lecture delivered to Centro de Estudios de las Tradiciones (El Colegio de Michoacán) graduate course Textos Amerindios I (Dr. Hans Roskamp). Zamora, October 17.

Advising
Englehardt, Joshua D.

Member of B.A. Thesis Supervisory Committees
Julio de Jesús García Fernández, Spring 2016. Propuesta de catálogo digital de la sala Olmeca del Museo de Antropología de Xalapa-MAX. Facultad de Antropología, Universidad Veracruzana. (Trabajo profesional de grado).

Slice, Dennis E.

Director of M.Sc. Thesis Supervisory Committees
Cameron Berkley, Spring 2016. Title TBD. Department of Scientific Computing, Florida State University.

Collaborations
We have formed collaborations with others who are working on project which could benefit either from the data structures that we have created and/or that are interested in integrating RTI and 3D imagery into their digital collections. One is with Mark Bonta of Pennsylvania State
University, whose work focuses on the ethnobotany of New World cycads, particularly those of Mexico. Bonta, Carrasco, and a team of other scholars from Mexico applied for an NEH Collaborative Research Grant that would have supported the extension of this project to cover botanical information. Recently, co-directors Englehardt and Carrasco have expanded their collaboration with Mexican ethnobotanical researchers, and have recently applied for a Frontiers of Science grant from CONACYT, the Mexican National Council of Science and Technology. The second collaboration is with the Florida Department of State and focuses on colonial period objects and earlier Weedon Island style ceramics that present iconographic issues similar to those encountered in the Formative period Mesoamerican corpus (see Appendix 17). Finally, co-director Englehardt, in conjunction with Verenice Heredia of the Proyecto Arqueológico Teuchitlán and the Ministry of Culture of the State of Jalisco, Mexico has applied the imaging techniques utilized in this project to the collections of the archaeological site of Los Guachimontones (see Appendix 18). The Ministry of Culture has expressed an interest in utilizing project data structures to catalog state collections, and has also asked for our assistance in creating a mobile application that can be incorporated into museum displays, with a design and digital tools similar to those in development through this project.

**Final Remarks**

The collection and preparation of data and the production of the digital architecture and assets have been completed as scheduled. We have made significant progress, and our accomplishments meet our established goals. New aspects of the project evolved with the introduction of the RTI and 3D Photogrammetry that allowed us to collect data that we had not originally intended to given the high costs associated with other forms of 3D scanning. These technologies have allowed us to widen the scope of the project and enrich its utility for researchers. RTI and 3D images have also permitted more accurate drawings and we have found are critical for assessing a monument’s current condition compared to historic
photographs. Our initial comparison of the photographic record with our new data reveals alarming levels of erosion and deterioration. Our work raises questions not only about the immediate need for greater protection of Olmec sculpture, but how best to present these objects when different moments of their documentational history differ. The possibility of digital reconstruction and the ethical and scientific issues related to this we hope to explore in a session we have organized for the 2016 World Archaeological Congress. In the end while aspects of the project have changed to accommodate new techniques of documentation, we have met all of the project goals set out in our original proposal during the grant period and exceeded our initial expectations for the documentation of objects.

Finally, while not our intention to generalize the digital architecture and applications from this project to other projects at this early date, we have nevertheless found potential future collaborators who are intrigued by the way we wish to organize data and would like to utilize various aspect of it in their own projects or in collaborative ones. We look forward to presenting these possibilities in more detail in our proposal for the Digital Humanities Implementation Grant.

We envision the educational reach of *The Mesoamerican Corpus of Formative Period Art and Writing* surpassing the presentation of essays and static visual data to allow for the dynamic search, visualization, and investigation of a corpus of material currently only partially available in scattered sources with restricted access. We see this project as a test run for the presentation of iconographic, contextual, and archaeological data serving many types of users whose structure has the potential to be used in humanities research that focuses on visual materials. By taking advantage of new digital technologies, we will create an innovative resource that will aggregate the various digital assets and tools discussed above. In this way we will move beyond a “website” or a searchable database of materials to produce a set of resources and applications that do not exist purely as a presentational medium, but rather allow for the active manipulation of data that leads to new ways of thinking about that material.
References

Bataineh, Bilal, Siti Norul Huda Sheikh Abdullah, and Khairuddin Omar

Batuwita, Rukshan, Vasile Palade, and Dharmapiya C. Bandara

Belongie, Serge, Jitendra Malik, and Jan Puzicha

Benson, Elizabeth P. and Beatrix de la Fuente (eds.)

Carrasco, Michael D.

Carrasco, Michael D. and Joshua D. Englehardt


Chavan, Urvashi R., and N. M. Shahane


Clark, John E. and Mary E. Pye (eds.)


Coe, Michael. D.


Dalitz, Christoph, Georgios K. Michalakis, and Christine Pranzas


Englehardt, Joshua D.


Frauel, Yann, Octavio Quesada, and Ernesto Bribiesca


Houston, Stephen D.


Hu, Rui, and John Collomosse


Hyder, Mashud, M D Monirul Islam, M A H Akhand, and Kazuyuki Murase


Joralemon, Peter D.

Justeson, John S.


Marshall, A. Malcom, and S. Gunasekaran


Pohl, Mary E. D., Kevin O. Pope, and Christopher von Nagy


Pohl, Mary E. D., J. Kathryn Josserand, Kevin O. Pope, and Christopher von Nagy


Quirarte, Jacinto

Reilly, F. Kent III


Rodríguez Martínez, María del Carmen, Ponciano Ortíz Ceballos, Michael D. Coe, Richard A. Diehl, Stephen D. Houston, Karl A. Taube, and Alfredo Delgado Calderón


Roman–Rangel, Edgar, Carlos Pallan, Jean–Marc Odobez, and Daniel Gatica–Perez


2011 Analyzing Ancient Maya Glyph Collections with Contextual Shape Descriptors.

Roman-Rangel, Edgar, Jean-Marc Odobez, and Daniel Gatica-Perez


Salton, Gerard, and Michael J. McGill


Schamndt–Besserat, Denise

2007 *When Writing Met Art: From Symbol to Story*. University of Texas Press, Austin.

2010 *How Writing Came About*. University of Texas Press, Austin.

Smith, Ray


Taube, Karl A.

Appendices

Appendix 1: Previously undocumented Middle Formative period celt from Río Pesquero. The right side of the image shows the object in way equivalent to a conventional digital photograph, though still from and RTI file. The left uses RTI technology to reveal details through raking light not possible in conventional photographs.
Appendix 2: A 3D rendering of a previously undocumented Early Formative Period Ceramic bust from San Lorenzo Tenochtitlán created using 3D photogrammetry technology.
Appendix 3: 3D renderings of the Early Formative Period stone figure known as the Señor de Las Limas. This page: frontal view; following page: top view and detail of were-jaguar infant cradled in arms of central figure.
Appendix 4: A screen shot of a 3D rendering of La Venta Altar 5 created using 3D photogrammetry technology.
Appendix 5. RTI (left) and project drawing of the Cascajal Block, the earliest known text in the New World. New visualization technologies allow for truer drawings, as well as the documentation of details such as tool marks and evidence of surface preparation, as seen in the detail image of the central throne-mat kenning (bottom).
Appendix 6: La Venta Altar 3, right side. This RTI image is the basis for our project drawing. To our knowledge, a drawing of this object has not been published. The use of new imaging technologies allows for the manipulation of the image, light raking, and vision from several angles. Drawings may then be produced using multiple perspectives, thus creating sketches of iconography that are more faithful to the original artistic program.
Appendix 7: (Top) An RTI (right side is enhanced) image of the front of La Venta Altar 4, which reveals details not often captured in conventional photographs or drawings; (Bottom) a screen capture of the 3D model of La Venta Altar 4.
Appendix 8: A still image taken from a 3D Photogrammetric model of La Venta Altar 4. Using this imaging technique, a face (outlined in box) and object (immediately to the right) have been documented on the right side of Altar 4. Although still not easy to see in the dynamic version of this 3D model of Altar 4 the face and object are more apparent.
Appendix 9: (Top) a historical photograph from the Smithsonian’s Stirling Collection, in which the face on the right side of La Venta Altar 4 is perhaps more visible. Nonetheless, new imaging techniques utilized in this project allowed for the production of a new drawing of the monument’s right side (below), thus recapturing details lost due to erosion.
Appendix 10: An RTI image of La Venta Monument 13 (right) compared with a historic photograph taken soon after its discovery. Notice the toes of the foot glyph to the figure’s left are no longer visible. Also, in the RTI photograph notice that the pockmarks which are not nearly as present in the earlier view. Over the last sixty years intrusions of softer stone have erode and left these marks. This appears to be a major issue with La Venta’s monuments. Overlaying RTI and light raking perspectives over historical photographs and previous sketches during the drawing process further enhances the fidelity and accuracy of project drawings (bottom: the preliminary project drawing of La Venta Monument 13).
Appendix 11 (following page). Preliminary UML map detailing planned digital architecture and relationships between system components.

- **Backend Interface**
  - Database
  - Script actions that move data from DB to user interface
  - Frontend User Interface(s)

- **Backend Data Input:**
  - This area is password protected but accessible to multiple users, particularly those in the field. There is a hierarchy of accessibility and editorial steps prior to data being integrated into the database.

- **User Input**
  - In working version registered users will be able to add to the database by completing the db fields. This will give the format the flexibility to deal with visual information from any region and coordinate this data with archaeological context.
  - A user manual will describe how to prepare the data and files to be entered.

- **Database**
  - Script to feed geo info to map interface and integrate with associate map layers (some native to the database others from third party).
  - Script to feed project imagery to specific viewers and locations in the user frontend interface.
  - Script to feed RTI files to viewer (the viewer needs to be developed).
  - Script to feed 3D models to MeshLab

- **Digital Tools native to Database**
  - Reverse image search
  - Geographical contextualization
  - RTI, PTM, 3D model viewer
  - Edge detection tool for working with images
  - Image layer viewer to dynamic use of the layered image files.

- **Frontend User Interface(s)**
  - Presentation across multiple devices
  - The interface presents an introductory page, Object view, site view with archaeological data, interactive map (leads to site view or object view), and motif catalogue.
  - Photo/Drawings, 3D, and RTI, viewers open to full screen mode.
Appendix 12 (following page). More detailed UMP map specifying information paths among system components and data structures, as well as relationships between informational items, database fields, and individual pages or views in the website and mobile application.
Appendix 13. Screenshot of map view with object information and iconographic data documents (on left) from prototype web application, displaying linkage between the MongoDB database to a map interface using OpenLayers. Compare against mockup screenshots in original proposal.
Appendix 14. Screenshot of site view with object information and iconographic data documents (on left) from prototype web application, displaying linkage between the MongoDB database to a map interface with GIS overlays using OpenLayers. Compare against mockup screenshots in original proposal.
Appendix 15. Screenshot of informational input fields from prototype website database, displaying JSON object IDs.
Appendix 16 (following page). Screenshot of informational fields from prototype website database.

<table>
<thead>
<tr>
<th>Item</th>
<th>Order</th>
<th>Name</th>
<th>Local Name</th>
<th>Institutional Catalogue</th>
<th>Catalog Number</th>
<th>Material</th>
<th>Dimensions</th>
<th>Weight</th>
<th>Site Known</th>
<th>Region</th>
<th>In-Site Location</th>
<th>In-Site Location</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>item1</td>
<td>item2</td>
<td>item3</td>
<td>item4</td>
<td>item5</td>
<td>item6</td>
<td>item7</td>
<td>item8</td>
<td>item9</td>
<td>item10</td>
<td>item11</td>
<td>item12</td>
</tr>
</tbody>
</table>

Additional table data follows.
Appendix 17: A 3D photogrammetric model of a colonial period sword hilt in the collections of the State of Florida Bureau of Archaeological Research. This image was produced by a participant in co-director Carrasco’s 2015 summer course in conjunction with the Florida Department of State Collections.
Appendix 18: 3D photogrammetric models of a ceramic figurine (top) and bowl of the Late Formative western Mexican Teuchitlán tradition from the archaeological collections of the site of Los Guachimontones, Jalisco, Mexico. These images were produced by co-director Englehardt in conjunction with the Proyecto Arqueológico Teuchitlán and the Secretaría de Cultura of Jalisco State.