

DATA MANAGEMENT PLAN

DIGITAL DIALECTIC: FORGING NEW PATHS OF INQUIRY IN THE HUMANITIES

Data Collection: We will be digitizing the fresco, *Mundos de Mestizaje*, using high resolution, high dynamic range, panoramic digital photography on site at the National Hispanic Cultural Center (NHCC). The images will be transported to the University of New Mexico where hundreds of digital images will be stitched together to create a composite panoramic image. Web site development and testing will be conducted at UNM. The interactively accessible embedded information and other curriculum content will be created at the NHCC concurrently with the technical development of software at UNM's ARTS Lab. Videos of user experiences in the fulldome, machinima of the web experience, as well as usability data will be collected at UNM. Documenting the tool building, production and use will be an ongoing process and we will collaborate with Data Librarians at UNM to accomplish this. Much of this will be accomplished by maintaining a project wiki/website.

Data Management: As high resolution photographs of the fresco are taken at the NHCC, they will be duplicated to two external hard drives (from the camera's compact flash card.) Upon completion of a day's shooting, the photos will be duplicated to a Level 5 RAID at ARTS Lab. The interactive metadata and content being developed will be stored locally and sent to a Level 5 RAID server at UNM via SFTP. A duplicate will be kept on a second RAID. The storage is supported with incremental backup to protect against data corruption or loss. All completed curricula and educational content will be kept in secure storage at UNM with back up along with the rest of the data. The National Hispanic Cultural Center is a division of the Department of Cultural Affairs (DCA) for the State of New Mexico. Websites are managed centrally by the department's Information and Technology staff. All website content, along with redundant copies of DCA data, is backed-up three times a week. The State of New Mexico's data center is the charge of the Department of Information and Technology and is housed off-site. As long as DCA hosts the website through the web-server the site will be available through the National Hispanic Cultural Center and Department of Cultural Affairs websites.

Data Types and Metadata: The original images will be high resolution Canon Raw (.cr2) format files. Post-processed images (HDRI and panoramic) will be appropriate 32-bit file formats supporting HDR; for example .hdr, .tiff or .fjpeg file. The software will be typical text files in Java and/or HTML5 languages with additional XML files associated. Our interactive content will be kept in HTML, jpeg/png for still images, .mp4/.mov for video and .ac3/.mpg2 for audio files. Machinima and videos of demonstrations will be stored in standard movie file types depending on the hardware and software used (e.g., .mp4 or .mov). In addition to the standard archival metadata outlined below, we will be working closely with Research Data Librarians at UNM to curate data. Technical metadata will be structured using MIX for image data and PBCore for digital movies. Additional metadata about the fresco will be supplied with VRA about the curriculum and lesson plans by IEEE LOM. The software will be documented appropriately in the code itself and any dependencies will be linked to the metadata records using RDF. These strategies for curating and archiving our data follow the guidelines outlined in <https://www.ideals.illinois.edu/handle/2142/17097> for preserving virtual worlds, which closely approximate the nature of this project.

Copyright and Licensing: Images derived from the fresco will be kept under full copyright, all rights reserved, although the images will still be viable on the web. Any additional images or non-software creative content, as well as the interactive metadata developed at the ARTS Lab will be released under a [Creative Commons Attribution-NoDerivs 3.0 Unported License](https://creativecommons.org/licenses/by-nd/3.0/). Software created for this project is intended to be released with an open source Apache License 2.0 <http://www.apache.org/licenses/LICENSE-2.0.html>.

Privacy and Sensitive Data Issues: On-line surveys of the web application's beta launch will not collect names, but will collect very basic social information (age, gender, education and regional location.) IP addresses will not be associated with the forms. Volunteers testing the software and previewing the fulldome content will also complete nameless feedback forms. No sensitive data is needed nor desired as part of the evaluation process.

Level of Data Shared: ARTS Lab will maintain a wiki dedicated to the project that will be made public at the beta phase. The wiki will include documentation of the photographic and software development processes, the generalized results of evaluations, and the source code for the web app. The wiki will be featured on both ARTS Lab's and NHCC's web sites. The published web site will include simple HTML access to the supporting content. The fulldome content will be made available via an ARTS Lab SFTP server upon request - for downloading and installation in participating theaters.

Archiving and Dissemination: This project incorporates data and resources from a variety of sources. Wherever possible we will make the data from this project freely available in our institutional repository or through other suitable outlets as described below. Where it is not possible to make a resource available, we will document the origin of the resource and the nature of the dependency and the terms of use in the metadata. Computer source code and other resources required by the website developed by this project will be made available through the project wiki (using Doxygen to document the code <http://www.stack.nl/~dimitri/doxygen/>) and made available using an SVN such as GitHub (<https://github.com/>) except where prohibited by other licensing agreements. Curricula, lesson plan, other unforeseen data, or links to data archived elsewhere, will be archived in perpetuity at the University of New Mexico repository. The data will be available no later than the conclusion of the grant in accordance to the rights policies outlined above. The data archive at the University of New Mexico is still being developed, but is being built using the Open Archive Information Systems (OAIS) model as its organizing paradigm. The DSpace repository uses Qualified Dublin Core for descriptive metadata. The data archive will keep separate metadata records in XFDU or RDF (linked data) architecture. We will use PREMIS as the primary administrative metadata schema. Additional technical and domain specific metadata schemes will be incorporated into the record in accordance to current standards in the field and as outlined above. The PI will be responsible for retaining the required metadata in consultation with the Research Data Librarians and work with the Data Librarians to organize and format the metadata, and ensure its completeness and accuracy.