

DATA MANAGEMENT PLAN: Onondaga Lake: Finding a restorative center in digital space

1 Roles and responsibilities

The project team, led by Rachel May and including Jane Read and Phil Arnold are all members of Syracuse University. Together they will be responsible for executing this project management and retention plan, with Project Director May overseeing implementation and coordination of the team's efforts. In the event of leadership change, Dr. Read will assume this responsibility. Each team member will be responsible for management of data specific to their area of expertise, e.g., Read: creation and evaluation of GIS and related mapping-related data; Arnold: conversations and artefacts associated with the Native Peoples, and May: data and artefacts associated with environmental and local knowledge. All team members will be responsible for reports and any articles resulting from the project.

2 Expected data

As this is a planning project, data will include a variety of types resulting from the developmental activities. These will include: audio recordings and written transcripts from interviews; geospatial data derived from historical and current maps from archives, governmental and non-governmental agencies, and other verifiable sources; geospatial data derived from results of focus groups and interviews; photographs and other images from archives and personal collections; and historical documents from archives and personal collections.

Due diligence searches of data.gov and other sites will be conducted prior to developing new datasets. Any data with personal identifiable information will be scrubbed of identifying information and the original files deleted permanently. Spatial data that contain personal identifiable location information will not be shared publicly and will be aggregated to the census block level or coarser.

Data will be stored on secure password-protected servers that are backed up nightly by the Maxwell School's Information, Computing and Technology group. All servers have a 30-day retention policy.

All datasets and reports will be preserved on the servers and shared with permission of the originators (e.g., focus group participants will be asked to sign an information release), with the exception of sensitive Native American information, any information with personal identifiable markers, or any other information that we consider might cause harm to an individual or group. Appropriate datasets will be deposited with the Syracuse University Maxwell School's Qualitative Data Repository (QDR), following their protocols on metadata and sharing while complying with copyright laws.

3 Period of data retention

All datasets (with their metadata) that are incorporated into the prototype, the online prototype pages, and the white paper will be deposited for indefinite storage with Syracuse University Maxwell School's Qualitative Data Repository (QDR - see attached letter of commitment). Aggregated, non-sensitive, and non-personal data will be made publicly available as soon as the prototype is completed and the white paper and any journal article about this first phase of our project is in press.

4 Data formats and dissemination

File formats will vary with content and method, but will include widely-adopted formats that can be shared and converted easily (e.g., Microsoft Word and PowerPoint, PDF, MP4, JPEG, TIFF, XML etc). Spatial data will be converted from ESRI proprietary datasets and shared as Keyhole Markup Language (KML) files. KML is an official Open Geospatial Consortium (OGC) standard and is commonly-used for sharing geographic data with non-GIS users as it can be easily delivered on the Internet and viewed in a number of free applications, including Google Earth and ArcGIS Explorer. We will consult with Syracuse University's QDR on the most appropriate file sharing formats for text, video and image files, in addition to ensuring that online file formats will be freely accessible, and the prototype itself, responsive.

All spatial datasets will be created following the Federal Geographic Data Committee's (FGDC) metadata standards. ESRI's ArcGIS software will allow us to view and edit metadata following the FGDC Content Standard for Digital Geospatial Metadata (CSDGM) guidelines, export metadata in this standard's XML format, and validate it using the CSDGM XML DTD. We will consult with Syracuse University's QDR on the most appropriate methods for documenting metadata for non-spatial datasets of varying formats (text, video, image etc).

The end goal of this project is to share data about Onondaga Lake publicly, and so we plan to share datasets online as soon as possible. Data and reports will be archived with Syracuse University's QDR as soon as the prototype is completed and the white paper and any journal article about this first phase of our project is in press. We anticipate that all data to be shared will be made public within two years of the end of the grant, but that many datasets will be made public sooner.

Prior to beginning the research, the Project Director will ensure that we have IRB approval, and will ensure appropriate protocols are followed. Any sensitive data that should be viewed only by certain parties (e.g., sensitive Native American knowledge) will be made available with permission only using password protected areas within the public site.

5 Data storage and preservation of access

All datasets (with their metadata) that are incorporated into the prototype, the online prototype pages, and the white paper will be archived and made available for long-term sharing at the Qualitative data Repository (QDR), where they will be curated and stored according to best practices for data management. QDR is hosted by the Syracuse University Maxwell School's Center for Qualitative and Multi-Method Inquiry. QDR is a dedicated archive for storing and sharing digital data (and accompanying documentation) generated or collected through qualitative and multi-method research in the social sciences (see attached letter of commitment).