Authors: David T. Hansen, Barbara D. Simpson, Anastasia Leigh, Patrick Welch, and Lorri Peltz-Lewis

# Geospatial Data Model for Archaeology Site Data

Presented by David T. Hansen and Barbara D. Simpson at the ESRI User Conference, 2005, San Diego California, July 28, 2005

Abstract

Description of a data model for managing archaeology data to address the needs of U.S. Bureau of Reclamation archaeologists. It has application for general model development for other archaeology or historic data. The Mid-Pacific Region of the U.S. Bureau of Reclamation (Reclamation) has collected a variety of archaeology information over the years to address issues related to the protection and conservation of this resource. Staff archaeologists would like to have this data in a readily accessible format for analysis in addressing impacts of agency actions to the resource. This data is from a variety of sources including State Historic Preservation Offices, other agencies, and internal site investigations. Organized by Reclamation project, the data has been collected and verified over time with formats and content of the data varying based on sources and project needs. This information includes the spatial representation of sites along with site records or field descriptions, images and photos. There is a need to bring this data together into a common structure that can meet the general requirements for the archeologists and fit into other enterprise efforts for the organization. The spatial representation of the sites includes points, linear transects or features, and polygons or area features. The geodatabase is designed to handle the variety of naming conventions and site attributes in use by the different organizations. This design will allow links to stored site records or images of sites. Site information stored in the geodatabase is available for use with statistical and geostatistical tools.

Introduction

For over ten years the Mid-Pacific Regional Office of the U.S. Bureau of Reclamation (Reclamation) has been in the process of capturing archaeological site data in GIS. This data has been collected over time to address issues for particular Reclamation projects and for analysis. Figure 1 shows the Mid-Pacific Region with a portion of the archaeological site information collected from various sources by staff archaeologists.
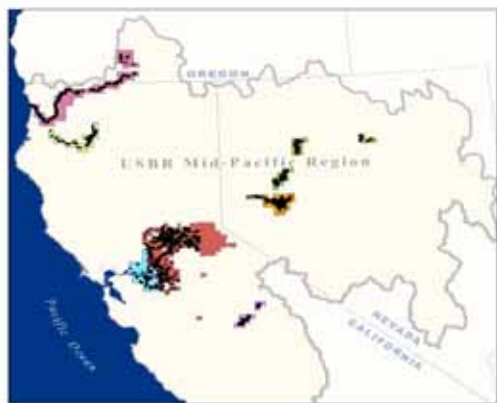
A significant portion of archaeological site information for the regional office is in GIS. This information has been collected from a variety of sources. These include various Offices of Historic Preservation (OHP), other agencies, and internal site inventories and records. Archaeological site information is considered by the Mid-Pacific Region to be sensitive data. It falls under a classification of "Sensitive Data" as identified in

Reclamation Memoranda on the handling and safeguarding of informaton and records. These are *Policy Memorandum - Interim Requirements and Procedures for Handling and Safeguarding the Bureau of Reclamation's Information and Records* and *Guidance Regarding the Handling of Information Relative to Entities with a "Need to Know"* (U.S. Bureau of Reclamation Memoranda, 2002). In the Mid-Pacific Region, staff archaeologists must review and approve access to and use of site information. This paper does not discuss issues of user identification and password control for access to the data or data encryption for storage or transfer. A geodatabase model has been under development to effectively manage this sensitive data with the following requirements:

- Identify the location of sites for the Mid-Pacific Region.
- Identify internal databases and database sources of site data.
- Link detailed site records and other documents to site locations.
- Standardize definitions and domains of valid values for reporting.

Geodatabase Model for Management of Archaeological Site Data

Conceptual Model

The intent of this geodatabase is to serve as a management tool for archaeology data. It is not intended to store all the information associated with a site. It identifies site locations and other databases of site information to assist in maintaining control of sensitive data for the Mid-Pacific Region. Definitions and terminology follow those used by various Offices of Historic Preservation (OHP) and the National Park Service (NPS). NPS is the source of standards and guidelines (NPS, Standards) which are implemented by the various OHP's. NPS maintains Web links to these State OHP's. NPS maintains the National Register of Historic Places (NPS, NR) and the National Register Information System (NRIS) (NPS, NRIS). Separate states maintain databases of site information. Requirements for the California information system are available from the California Office of Historic Preservation (CaOHP, 2002) Web site. In addition, NPS has developed a comprehensive archaeological sites management information system (ASMIS) (NPS, ASMIS). NPS serves as the lead agency for the Department of Interior for identifying requirements of Federal law and regulations governing cultural resources.

Identification of site locations- The geodatabase stores a site location for each recorded site. Coordinate values for site locations are based on the geodatabase spatial domain. Any other coordinate values for site location are stored independently of the geodatabase. Topology rules are set to avoid duplicate representations of a site.

Identify internal databases- Independent databases or files associated with a site are identified. This provides access to additional information for a site not contained in the geodatabase. It also assists in managing and controlling access to site information. A relationship is established between site identifiers or names for access to the independent databases to handle variations in definitions and formats.

Linkage to site records and other documents- Staff archeologist prefer access to the detailed site record, reports, or source maps where they are available. The database permits the linking of site records and related documents to site locations.

Standardize definitions and domains of valid values- The variety of field definitions and domains of valid codes used over time have hindered the application and use of site information. This model relies on many of the definitions and domains of valid codes used in ASMIS. It does not duplicate the full structure of ASMIS. Portions of the ASMIS domains are carried that have immediate application for the data management needs of Reclamation archaeologists. Additional fields and domains are added to meet specific requirements for Reclamation. Portions follow other Federal Geographic Data Committee (FGDC) standards

for geospatial data. Typically, these portions are related to the reporting of coordinate locations or assessments of spatial accuracy.

The geodatabase consists of a set of feature classes, related tables with defined relationships, and domains for controlling definitions and vocabulary. Figure 2 shows the overall structure of the site management geodatabase.
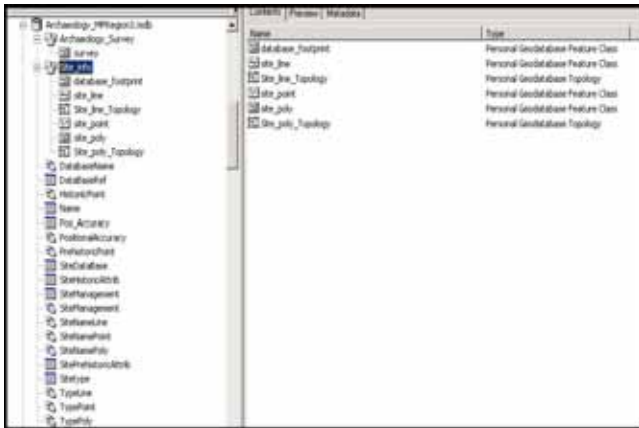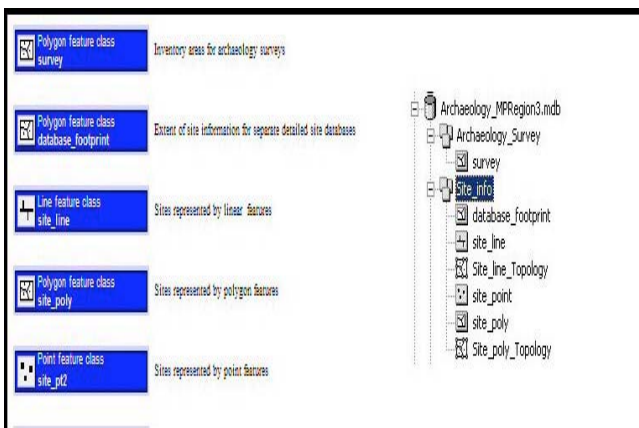

Feature Classes

There are a set of feature classes within two feature datasets. These are:

- Survey - A polygon feature class representing areas of archaeology surveys.
- Database footprint - A polygon feature class identifying the extent of independent databases of site information.
- Site_Line - A line feature class representing linear archaeology sites.
- Site_Point - A point feature class representing archaeology sites as points.
- Site_Poly - A polygon feature class representing archaeology sites as polygons.

The survey feature class is used to track the area of any completed surveys. The feature classes containing site locations are the main set of feature classes for accessing the site location information. The databaset_footprint feature class is used to identify the extent of independent databases that may contain additional site information. Figure 3 shows the set of feature classes within the feature datasets.

Site locations (points, polygons, lines)- These feature classes represent the main component of the geodatabase. As a management database, a site is represented only once. The geodatabase model assists in preventing duplicate loading of a site multiple times from separate project databases that have developed over time. Where point locations are needed for analysis of site distribution, polygons or lines are exported

as point data.

The structure of these feature class tables are kept simple. Key fields of SiteID and Name are maintained for relating to other tables. Relationship classes are established on these fields for linking to the other tables in the database.

The point feature class contains additional fields for storing coordinate information, surface elevation and reported elevation for the site. The coordinate information is based on the spatial domain of the feature dataset. Locations originally reported as part of the site information are stored only with the source data. This includes commonly reported UTM coordinate locations as well as the location within the Public Land Survey System (PLSS). Where PLSS is needed for reporting, it is generated based on a separate database of PLSS following the National Integrated Land System (NILS) formats (citation). Figure 4 shows an area with site locations and the attribute information carried for the point data.

Foot prints of source databases- This feature class contains the extent of independent databases of site information. The foot prints may represent other geodatabases, shape files, ArcInfo coverages. They could represent files with coordinate values of site locations. This feature class identifies any independent databases with site information. It assists in managing and controlling sensitive data. The feature class provides a spatial reference to these independent databases. The databases may contain more site information than the information contained in the geodatabase tables. Database foot prints are shown in Figure 1 for the Mid-Pacific Region.

Related Tables

Associated tables carry most of the attribute information for the sites. They also identify databases with additional site information or links to site records and documents. The set of tables is expected to be dynamic. Additional tables will be added as needed for management of the site information. Currently, main tables are:

Name
> This table carries names or identifiers used for the site. Often there are at least two different names associated with a site. Each site may carry an initial temporary number and assigned number from the local OHP. This assigned number is often referred to as the SHPO number or the Smithsonian number. In addition, there may be an assigned Primary number by the OHP. It is similar to tables ASMIS and AltDesig in the NPS database.

Prehistoric Attributes
> Prehistoric attributes associated with the site based on a defined domain. One site may have many associated prehistoric attributes associated with the site. The domain is similar to the NPS site and feature type.

Historic Attributes

    Historic attributes associated with the site based on a defined domain. One site may have many associated prehistoric attributes associated with the site. The domain is similar to the NPS site and feature type.

Site Type

    This table identifies main characteristics for the site including time period represented, presence of human remains, and site conditions. There is a one to one relationship between a site name and the records in this table. Domains for time period and site condition are similar to the domains carried in the NPS database.

Site Management

    This table identifies the agencies and organizations having jurisdictional and administrative responsibilities for the site, National Register status, and collection status for the site.
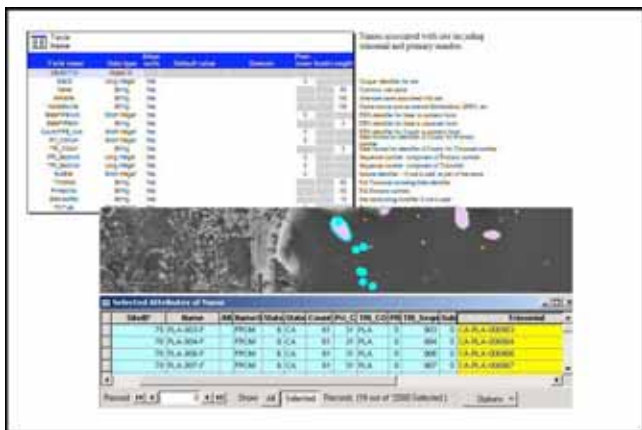
Positional Accuracy

    This table identifies the positional accuracy associated with the digital representation of the site location. A key component associated with location information is the accuracy associated with the sites based on the coordinate source. The domain for the reported source of the coordinate values follows the domain used in the NPS ASMIS database. Accuracy assessment follows the requirements of the FGDC content standards.

Independent Databases

    This table identifies the source database for the site or the link to the site record. Many of the sites are from sources external to Reclamation. The associated database carries information in the original format for these sites. A separate table describes the source database and serves as a reference to the storage location and status of the database. It serves as a control for these site records.

Name Table - The most complex table in this group is the Name table. Variations are common for names and other site identifiers. In the inventory process, an initial site identifier or temporary number is assigned. This will be replaced by an identifier assigned by the State Historic Preservation Officer (SHPO). These identifiers are frequently referred to as trinomials (also referred to as Smithsonian site number). In some areas, the trinomials are being supplemented by a primary number. Besides carrying a common name of interest to Reclamation staff archaeologists, the table carries names for the site from the source and components for constructing a formal Trinomial and a Primary number. Figure 5 shows the separate fields carried for this table. As site information is reported to the National Register of NPS, an additional unique identifier may be assigned.


The primary fields for this table are SiteID and Name which may be used to relate to other tables within the database.

Prehistoric Attribute and Historic Attribute Tables - In comparison to the name table, the tables for identifying historic or prehistoric attributes for a site are deceptively simple. They carry the types of features

identified at the site from a controlled set of vocabulary. A site may have many features reported. Feature types recorded in the table generally represent a subset captured from the site record. The potential domain of values for either prehistoric or historic feature types in very large. The domain actually used for the geodatabase is limited to a few key terms of interest to the staff archaeologists. These were selected after extensive review of feature types identified by the California OHP guide (CaOHP, 2002) and NPS ASMIS data dictionary (NPS, ASMIS, 2003). Categorization of feature types reported for a site is always open for discussion by archaeologists. The domain of defined site or feature types is not considered static. Access to the actual site record can assist in categorizing sites for analysis and in refining the domain of site types. Viewing the site record is often preferred by the archaeologists.
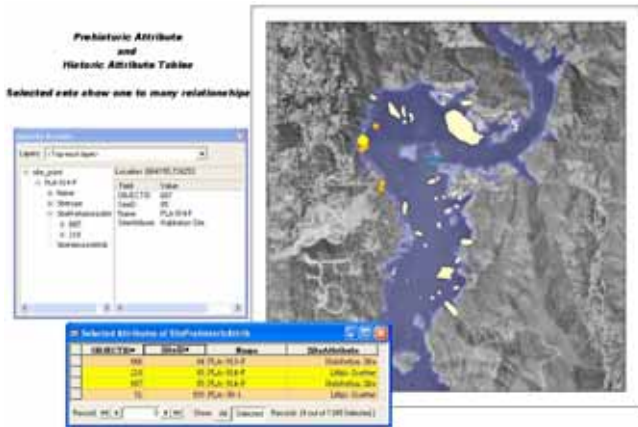
Domains

Domains assist in standardizing information associated with the sites. To the extent possible, these domains are based on definitions and the vocabulary in the NPS ASMIS database. ASMIS serves as an underlying template for:

- Referral to format, authorities, and domains of commonly used values.
- Source of additional table and field definitions for information that may be incorporated into the model at some future date.
- Relationship of model field definitions and domains to similar definitions in ASMIS as cross walks for reporting.

These and related terms from other authorities such as OHP's or FGDC standards have been reviewed by staff archeologists to identify those terms which are either necessary or most useful for the sites.

Domains include:

Time Period
> There are two domains for general time period associated with a site. Identification of sites dating prior to or after contact of Native Americans with foreign persons is of primary interest to staff archaeologists. Some sites contain features from both time periods. Also of interest to staff archaeologists is the domain defined in ASMIS as General Time Period. Figure 7 shows the domains for these general time period domains. Domains include the following values:
>
> - Unknown - Period of use is not known.
> - Prehistoric - Features date prior to Native American contact with foreign persons.
> - Protohistoric - Features date from initial Native American contact with foreign persons.
> - Historic - Features date after arrival of foreign persons
> - Modern - Features date from present time to 50 years ago.

Positional Source

The values in this domain identify the original source of the coordinate values for the site locations. Values carried in this domain are based on the set carried in the ASMIS database. They form the basis for assessing the accuracy in the coordinate values for the site location. Figure 8 shows the domain for positional source paired with a domain for positional accuracy. The domain includes:

- Feature compiled on a 7.5 min. USGS quad map.
- Feature captured from a historic map.
- Feature location based on reported location within cadastral survey such as the Federal Public Land Survey System (PLSS).
- Feature captured from aerial photography.
- Feature captured with resource grade or survey grade Global Positioning System (GPS).

Positional Accuracy

This is a defined set of accuracy values in meters. It carries appropriate values for positional accuracy based on the source of the coordinate values. It is based on the set of values identified in ASMIS for positional accuracy. This is also consistent with the positional accuracy assessment element of FGDC standards (citation). Figure 8 shows the domain for positional accuracy and associated domain for positional source.

Prehistoric Attributes

This is a set of terms defined by Reclamation staff archaeologists for representing the types of features found at a prehistoric site. It is based on a review of site and feature types defined in ASMIS as well as feature types identified by separate OHPs. Related to this domain is a separate field that identifies evidence of any human remains or burials. Feature types include:

- Isolate
- Lithic scatter
- Camp site
- Habitation site
- Rock art
- Bedrock milling station

Historic Attributes

This is a set of terms defined by Reclamation staff archaeologists for representing the types of features found at an historic site. It is based on a review of site and feature types defined in ASMIS as well as feature types identified by separate OHPs. Feature types include:

- Isolate
- Trash scatter
- Structure
- Surface modification

Jurisdiction

This consists of a set of controlled terms to identify the surface owner at the site location. It follows the set of values for jurisdiction identified in ASMIS. In the table, it is paired with a field that identifies the surface managing agency for lands under Federal jurisdiction. Figure 9 shows the domain for surface jurisdiction and in this case the paired field for managing agency. This domain includes:

- Federal land
- State government

- Local government
- Tribal
- Private
- Not determined

National Register Status

This is site status related to listing on the National Register of Historic Places. For the Reclamation database, it represents a partial set of the ASMIS domain. This domain includes:

- Recommended eligible
- Ineligible
- Listed
- Removed

Site Condition

This is an evaluation of the site condition as last reviewed by the archaeologist. It represents a subset of values from ASMIS for threat or disturbance categories. Related to this field are additional fields that identify sites which are inundated or are destroyed. Domain includes:

- Intact
- Disturbed
- Unknown

Collection status

This domain identifies archaeological work done at the site and it represents a subset of values from ASMIS identified as work assessment. For sites where collection has occurred, additional fields identify museum properties and name of the museum collection. The domain includes:

- Undetermined
- None
- Surface collection
- Tested (Partially excavated)
- Excavated / Mitigated
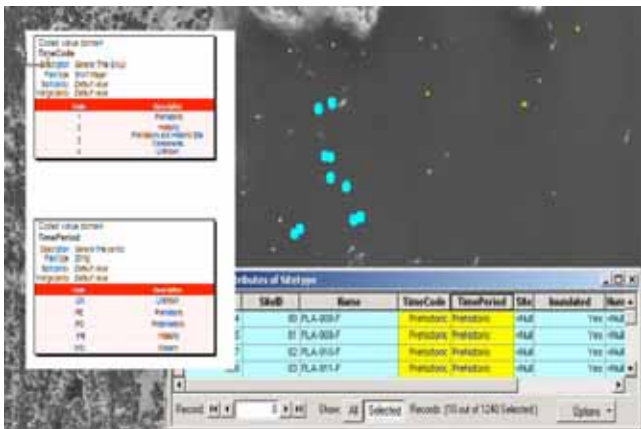
Summary

This geodatabase is intended to identify site locations, sources of site information, key attributes, and archeological site databases for the Mid-Pacific Region. It is intended as a management tool to handle and control sensitive archeological site data. It includes information to support reporting to the separate Offices of Historic Preservation and NPS. Much of the terminology and domains of acceptable values are based on those defined in the NPS ASMIS database.

Reclamation has begun the process of identifying key requirements and characteristics for a Reclamation wide database for archaeology site information. All Reclamation offices report site information to various Offices of Historic Preservation (OHP) and to the National Park Service (NPS). There are variations in the usage and application of terminology between these offices. It is expected that an overall model for archaeological site information will develop over time and that the various OHP's and the NPS will play major roles in the model development. Part of the effort in the development of this Reclamation geodatabase is the documentation of definitions and domains in current use. It is hoped that this model will serve as a basis for further discussion of a basic archaeology geodatabase model.

Acknowledgement

References

California Office of Historic Preservation, Department of Parks and Recreation, (CaOHP), *Information*

Center Procedural Manual, California Historical Resources Information System, California Resources Agency, November 20, 1995 as amended through 2002, Available via: http://ohp.parks.ca.gov/pages/1069/files/2001%201C%20Manual.pdf. http://ohp.parks.ca.gov/nr/index.html.

Federal Geographic Data Committee (FGDC), *Content Standards for Digital Geospatial Metadata* ,CSDGM Version 2 - FGDC-STD-001-1998 Washington D.C. 1998. http://www.fgdc.gov/metadata/contstan.html

National Park Service, Department of Interior (NPS, Standards) , *Archeology and Historic Presevation: Secretary of the interior's Standards and Guidelines*, May 25, 2005. http://www.cr.nps.gov/local-law/arch_stnds_0.htm.

National Park Service, Department of Interior, (NPS, NRIS), *National Register Information System (NRIS)*, August 1, 2004. http://www.cr.nps.gov/nr/index.html.

National Park Service, Department of Interior (NPS, NR), *National Register of Historic Places*, May 25, 2005. http://www.cr.nps.gov/nr/index.html.

National Park Service, Department of Interior (NPS, ASMIS, 2003), *Archeological Sites Management Information System Data Dictionary, version 2.03* December 2003, Archeology and Ethnography Program National Center for Cultural Resources, NPS, Washington D.C.

National Park Service, Department of Interior, (NPS, ASMIS, 2005), *Archeological Sites Management Information System User Guide, version 3.0* February 2005, Archeology and Ethnography Program National Center for Cultural Resources, NPS, Washington D.C.

U.S. Bureau of Land Management (BLM), *NILS Geodatabase Design - An ArcGIS Federal Lands Data Model*, ESRI Redlands CA, September 27 2002. http://support.esri.com/ under DataModels > Land Parcels.

U.S. Bureau of Reclamation, *Policy Memorandum - Interim Requirements and Procedures for Handling and Safeguarding the Bureau of Reclamation's Information and Records*, Memorandum by John W. keys III, Commissioner, signed June 19, 2002, D-5500, RIM-1.00.

U.S. Bureau of Reclamation, *Guidance Regarding the Handling of Information Relative to Entities with a "Need to Know"*, Memorandum by Larry L. Todd, Director of Security, Safety, and Law Enforcement, signed August 21, 2002, D-5500.

Authors:

David T. Hansen
GIS Specialist / Soil Scientist
MPGIS
Phone: (916) 978-5268
FAX: (916) 978-5290
Email: dhansen@mp.usbr.gov

Barbara D. Simpson

Gis Specialist
Division of Environmental Affairs
Phone: (916) 978-5031
Email: bsimpson@mp.usbr.gov

Anastasia Leigh
Archaeologist
Division of Environmental Affairs
Phone: (916) 978-5568
Email: aleigh@mp.usbr.gov

Patrick Welch
Archaeologist
Division of Environmental Affairs
Phone: (916) 978-5040
Email: pwelch@mp.usbr.gov

Lorri Peltz-Lewis
GIS Database Administrator
MPGIS
Phone: (916) 978-5271
Email: lpeltzlewis@mp.usbr.gov

U.S. Bureau of Reclamation
Mid-Pacific Region
2800 Cottage Way
Sacramento, CA. USA 95825-1898
Phone: (916) 978-5266
FAX: (916) 978-5290
Email: cbjohnson@mp.usbr.gov