

Colorado SHPO GIS Data Capture Procedures and System Design
2/28/01

Josh Torres

- All sites and isolated finds are represented as lines or polygons. Sites under 5 acres are the result of buffered points based on user specified tolerances representative of the size of the site.
- Sites over 5 acres will be digitized as they are represented on maps submitted to the SHPO. Sites over 5 acres, with poor spatial representation, will be handled like sites under 5 acres with the appropriate buffer tolerances applied to features based on the size of the site (Table 1).

Size of the Site (Acres)	Buffer Tolerance	Acreage of Resultant Polygon
IF	3	.007
>-.5	18	.25
.5-1	25.5	.502
1-3	36	1.001
3-5	62.5	3.017
5-10	82.5	5.001
10-15	114	10.037
15-20	139.5	15.030
20+	161	20.020

Table 1: Buffer tolerances for sites.

- Linear sites will also be represented as polygons (Although a polyline theme for sites will be generated). Linear features will be buffered to a specified user tolerance and merged to the polygon shapefile.
- Only the most recent iteration of sites will be recorded in the GIS.

Colorado SHPO GIS Digitization Standards for Surveys

- All surveys are represented as polygon features.
- Linear survey features are buffered lines to a given specified user tolerance.

The majority of the data has been digitized via the “heads up” method utilizing discrete topographical features to define the bounding coordinates of respective polygon features (typically at a scale of 1:5,000 in the view). Digital Raster Graphic maps (DRGs) that represent 7.5-minute USGS quadrangles were utilized as base maps for the digitization of survey areas and sites. The respective spatial features were digitized upon the DRG basemap sources and saved in ArcView’s native shapefile format. In doing so, it

has been possible to save time, decrease the introduction of error into the GIS database, and guarantee relatively accurate representation of areas evaluated for cultural resources.

As the spatial extent of archaeological sites are often difficult to quantify (c.f. Dunnell 1992:22-32), some methodological standards were developed in the capture of site data into a GIS. Polygon shapefiles representing sites are the result of buffered point locations for sites between 0 and 5 acres. Sites over 5 acres are digitized as they are represented on the site form maps submitted to the Colorado SHPO. Sites between 5 and 10 acres with poor polygon representation will be buffered to 80.5 meters (5.005 acres). Linear sites and survey areas were buffered to 15 meters (roughly 100 feet) unless otherwise specified. These buffered line and point features were then merged with the respective polygon survey or site data set.

Although well organized, the composition of the B. Rec. and BLM data is represented by hand drawn features on report documentation that display specific survey areas and site locations. Accuracy to which these project areas are representative of their “real world” counterparts is a function of how precise these areas were recorded by B. Rec. BLM personnel, and their contractors. Therefore, tablet digitization of these features from the reproduced maps was not only deemed impractical, in terms of time and effort expenditure, but was not necessary for accurately representing these features. To elucidate, there are three aspects relating to the data which clarify the decision to utilize the “heads up” method:

- (1) The distortion of the maps as a result of replication (i.e. photocopying).
- (2) The inability to accurately acquire “tics”, or referenceable coordinates, from these portions of replicated maps.
- (3) The reduction of time and effort expenditure through digitizing directly from the source maps.

However, in the event that locational data exists for any project or resource are on a stable medium efforts will be made to tablet digitize these features. With the use of these methods, in the appropriate contexts, it is possible to capture data within acceptable standards. *Given the numerous sources of potential error, data has been captured in the most efficient and precise manner possible to accurately represent survey and site features spatially.*

Attribute Definition Associated with the Spatial Data

Tabular data associated with the B. Rec. and BLM spatial data layers has been derived from report documentation and database files that are housed at the SHPO. The tabular data associated with the spatial data sets is standardized for polygon survey areas and site features. The base tabular data associated with the survey and site polygon shapefiles and an explanation of the attributes is listed in Tables 2 and 3.

Attribute	Definition
SHAPE	Shape of the spatial features in the data set. In this case polygons.
ID	Unique sequential numeric ID for a given

	spatial feature.
AREA	Area of the spatial features in the data set
PERIMETER	Perimeter of spatial features in the data set
ACRES	Acreage of the survey area calculated by the GIS from the spatial features in the data set
DOC#	Unique SHPO number referring to a specific report document.
AGENCY#	Unique agency identifier.
CONF	Confidence given to the spatial accuracy of the digitized feature. Values for this attribute consist of LC (Low Confidence) or HC (High Confidence).
VER	Verification. Refers to verification of the survey boundaries completed by the individual digitizing the spatial feature. Values for this field will consist of the initials of the individual who digitized the spatial feature.
Date	Date the feature was digitized.
Zone	UTM Zone encompassing the spatial feature.
X	The X coordinate of the center point of the survey.
Y	The Y coordinate of the center point of the survey.

Table 2: Naming convention and description of survey polygon shapefile attributes.

Attribute	Definition
SHAPE	Shape of the spatial features in the data set. In this case, these are polygons.
ID	Unique sequential numeric ID for a given spatial feature.
AREA	Area of the spatial features in the data set
PERIMETER/LENGTH*	Perimeter of spatial features in the data set. * This attribute is called Length for linear data layers.
ACRE	Acreage of the site calculated by the GIS from the spatial features in the data set
SITE#	Site number
BND_CMPLT	Boundary completeness. Refers to the completeness of the site boundary. Values for this field will either be Y (YES the boundary is complete) or N (NO the boundary of the site is not complete or unknown).

VER	Verification. Refers to verification of the site boundaries completed by the individual digitizing the spatial feature. Values for this field will consist of the initials of the individual who digitized the spatial feature.
Date	Date site was digitized.
Zone	UTM Zone encompassing the spatial feature.
X	The X coordinate of the center point of the site.
Y	The Y coordinate of the center point of the site.

Table 3: Naming convention and description of site polygon shapefile attributes associated with the spatial data.

Tabular Data

Although the base tabular information provides important data regarding the spatial features, this alone is insufficient for assessing projects and the cultural resources resulting from them. Therefore, accompanying the spatial data layers are separate tables that contain information regarding the geospatially-referenced survey areas and sites. These tables are compiled through file searches from a central database at the Colorado SHPO. The tables can then be “linked” or “joined” to the Doc# or Site# fields depending on whether the user is working with surveys or sites. Importantly, organization of the data in this manner allows for a dynamic method of cultural resource data management. Putatively, it makes the building of the thematic data layers (and ideally their maintenance) economic in terms of time, organization, and effort expenditure. Secondly, through the simplicity of the tabular data directly associated with the shapefiles, it is possible for any given agency to utilize the Agency# field as a linking field to data housed at the respective agency and (presumably referenceable) by their unique number. This allows the thematic layers more versatility and capability of customization based on the specific needs of the given agency. As the spatial data are updated to reflect the current standing of survey and site locations the associated tabular data regarding these spatial features can be re-queried and re-joined to the spatial data supplying an up to date and accurate representation of these features.

Relevant attributes regarding the surveys and sites were selected based on an inter-organization constituent user survey conducted by Colorado SHPO in July of 1998 and modified slightly upon commencement of this project. These attributes are described and listed in Table 4 and 5.

Attribute	Description
ID	Refers to the document number relevant to the project. This field will be used to “link” to the Doc# field in the survey polygon shapefile table.
Name	Name of the survey.
Completion.Date	Date the survey was completed.
Doc.Author	Author of the report resulting from the survey.

Method	Survey method used for the survey (e.g. class II or III).
Institution	Institution conducting the survey.
Site.Count	Number of sites resulting from the survey.
IF.Count	Number of Isolated finds resulting from the survey.
Acres.Total	Total number of acres <i>reported</i> for the survey.

Table 4: Naming convention and description of SHPO derived data attributes for polygon survey areas.

Attribute	Description
ID	The Smithsonian numeric site designation.
Site.Name	Name of the site.
Site.Type	Site type of the resource
Assessment	Assessment of eligibility for national register status.
Assessment.Date	Date of assessment for the resource
Culture	The cultural affiliation of the occupants of the site.
Feature	Any recorded features at the site.
Feature.Count	Number of features recorded at the site.
Artifact	Artifact types found at the site
Artifact.Count	Number of artifacts
Architecture.Early.Date	Early construction date associated with any historic architectural features at the site.
Architecture.Late.Date	Late date associated with any historic architectural features at the site.
Architecture.Site.Type	Historic site type
Architecture.Original.Use	Original use of the structure.
Architecture.Present.Use	Present use of the structure.
Architecture.Style	Style of architecture at the site
Recorder	Individual recording the site.
Organization	Organization that the individual recording the site is affiliated with.
East	UTM Easting/s of the site.
North	UTM Northing/s of the site.

Table 5: Naming convention and description of SHPO derived data attributes for site polygon features.

System Testing

Upon completion of a data set, quality control will be performed on the tabular and spatial data. Specifically, records will be double checked for the following items:

- (1) To ensure there is no orphan or missing data and that the tabular data has been entered properly.
- (2) To ensure the spatial integrity of represented survey and site areas based on comparisons with relevant project reports.
- (3) To ensure spatial integrity through scrutiny of the spatial data files.

Once the data has been thoroughly checked it will be disseminated to the appropriate agencies for review. Once at the appropriate agency, the shapefiles should be tested by the cultural resource staff for accuracy and functionality. If it is ascertained that the procedures need to be modified, discussions will take place to address these issues and their rectification.

- **Data Referencing and Documentation**

Shapefiles created for this project are documented utilizing FGDC metadata standard requirements. The metadata documentation follows the guidelines set forth by Gnomon Inc., the Wyoming SHPO and the Archaeological Records Management Section of the New Mexico SHPO which were submitted to the U.S. Geological Survey Federal Geographic Data Committee (Ingbar et al. 1999). Metadata for the datasets has been compiled using The Army Corps of Engineers metadata compiler "[Corpsmet95](#)".

- **Data Maintenance and Perpetuation of the SHPO GIS Program**

Ideally, the BLM data layers will be maintained by qualified personnel at the respective resource area field office. As the B. Rec. data layers represent "completed" mitigation projects they will not be updated. Currently, the Colorado SHPO is making steps towards creating statewide data sets for surveys and sites. The BLM and B. Rec. shapefiles will be incorporated into these "master" shapefiles. The Colorado SHPO is currently planning to initiate the digitization of all incoming material (i.e. survey and site data) into a GIS tentatively starting Nov. 2000. Cleanup of Colorado SHPO legacy data and the creation of datasets for the BLM and the B. Rec. will be ongoing and merged to the Colorado SHPO master datasets. The SHPO datasets will constantly be updated as new data is submitted.

- **Conclusion**

In conclusion, GIS utilized in cultural resource management contexts provides a new and powerful method for organizing and geospatially referencing data. Ideally, the Colorado SHPO should possess integrated thematic data layers which reflect *all* areas surveyed by various agencies for any given area or time period in the state. Further, data for sites should also be created and made available. Potentially, this will assist researchers and project managers in protecting Colorado's cultural resources. With the initialization of this program through the digital capturing of BLM and B. Rec. data, issues regarding data organization and methodological procedures will no doubt need to be modified to accommodate various constituents needs. Importantly, this program signifies a step towards the future for increasing the functionality and integration of Colorado's cultural resource database.

References

Dunnell, Robert C.

1992 The Notion of Site. In *Space, Time, and Archaeological Landscapes*, edited J. Rossignol and L. Wandsnider. Plenum Press, New York.

Ingbar, Eric E., Mary Hopkins and Timothy Seaman
1999 *Creating a Cultural Resources Metadata Standard for the Western United States*. A report submitted to the U.S. Geological Survey Federal Geographic Data Committee in partial fulfillment of Cooperative Agreement #1434HQ97AG01904.



METADATA CATALOG

Links to GIS information and data sources:

U.S. Geological Survey Data

<http://mapping.usgs.gov/nsdi/>

Bureau of Land Management GIS Data—Colorado Page

<http://www.co.blm.gov/metadata/cothemes.htm>

National Atlas of the United States

<http://www.usgs.gov/atlas/>

FGDC Standards

<http://www.fgdc.gov/standards/standards.html>

Standards for GIS Datasets in Cultural Resource Contexts

<http://colby.uwyo.edu/fgdcncptthome.html>