Introduction: The Need for a Standard

Standards for the content and format of data sets enhance efficiency in many ways. Standards increase the value of data by defining content, so data users know how best to employ information. Standards increase the exchange and needless duplication of information collection, by creating reliable formats for transmitting data. Standards define target formats for new software, hardware, and work processes, saving effort as an enterprise evolves.

This best practices memorandum provides formats in which spatial information in GIS data file formats should be conveyed to BLM information systems from other parties. BLM will use this same basic format to convey information to other parties, such as SHPO information systems. Although use of GPS is an important reason for this proposed standard, the standard itself does not require that GPS be used to create spatial information. GPS data transmission standards and mapping in general are addressed in a short appendix to this document (Appendix A).

The accuracy and reliability of a GIS entity (a point, a line, a polygon, a cell) are its most important spatial characteristics from an information user perspective. For most purposes, how something came to be in the GIS is less interesting than its meaning. Certainly, there are times when one might wish to know the signal strength of the specific GPS satellites used to determine a position – it's just that such times are rare for most users. The standards here are a template that will be used in every western U.S. state to convey GIS datasets from one organization to another.

The standard here pertains to conveying data as GIS data files. In general, we expect that organizations capable of creating GIS records are using GPS as a mapping tool. There is no necessary equivalence here: one could be using just GIS to create records. The standards are indifferent as to the source of spatial information. Instead the standards focus on reasonable description of the information's accuracy. Also, and importantly, the standard can be extended by adding columns to the attribute tables. So, this standard establishes a baseline format and values which *must* be followed, but allows for additional information to be built upon the same structure.

Datasets have value when the quality of information they contain is known and evident. *Metadata* describe the status of information within a dataset, documenting information quality. There are numerous discussions of metadata (see www.fgdc.gov for spatial metadata); a discussion of metadata in cultural resources is part of Appendix A. Most of the standard presented below is concerned with documenting the quality of information in a GIS; in other words, metadata. Four simple values are used to describe the spatial quality of cultural resource information.

A brief note on terminology is necessary. Throughout this document "phenomenon" and "entity" are used in very specific meanings. A *phenomenon* is the real world "thing" that is mapped or recorded. An *entity* is the representation of a phenomenon in an information system. Usually, *entity* refers to a GIS entity (a graphic representation in a known coordinate space), but it could also be a row of data in a table.

Best Practices and Implementation

This standards document grew out of a working collaboration between several different organizations. It establishes standards and is a starting point for anyone who needs to build a cultural resources GIS data file that will ultimately be transmitted to a BLM or a SHPO cultural resources information system. One can certainly add on to the information categories presented here. Like any standard, deviating from it or changing values in it will make it weaker.

There are two kinds of information about GIS datasets that are part of this standard: dataset-level metadata that describes the entire set of entities in a particular GIS dataset and entity-level metadata that describes each component entity. The two have different documentation methods. The former can be documented with tools available in current GIS software, the latter requires adding columns to specific GIS attribute tables (or other tables keyed to them).

BLM currently supports the ESRI ArcGIS data file formats. Portable (transmittable) formats are shapefiles and ESRI personal geodatabases. An additional advantage of the ESRI software is their ability to create overall dataset-level metadata readily. However, ArcGIS is only one of many GIS tools in common use. Almost all of them support the creation of dataset-level metadata with some form of tool. ArcView 3.x has its own set of add-in extensions for metadata, as does AutoCAD Map, and GeoMedia. In short, the software platform used for GIS is not a barrier to creating dataset-level metadata.

Entity-level data and metadata are not created by any of the applications described above. This is because the entity-level information is in specific columns in GIS attribute tables. These data and metadata columns, or fields, can be created in almost any GIS software as part of an entity-associated table. Models for these can be found in ESRI's personal geodatabase format, as Microsoft Access tables, and as ESRI shapefiles.

File Format Standards

Various file formats are appropriate for transmitting data to agency offices. The most useful are non-topological (so that GIS entities can overlap each other) and be easily read by various software tools. The ESRI shapefile format is perhaps the most universal format currently used for GIS. The shapefile format is **recommended** (but not mandatory) for transmitting spatial data with attributes at the present time.

Datum and Coordinate System Standards

The most important point about datum and coordinate systems is that they be well-documented for each dataset transmitted from one party to another. Many BLM administrative units have their own standards (usually at a state level). BLM CA employs the following standards:

- Datum: North American Datum of 1983 (NAD83)
- Coordinate System: Universal Transverse Mercator (UTM) in the appropriate zone (10 or 11)
- Datum Conversions (if needed): use HARN if available, NADCON if not

Data Set Level Metadata

Dataset metadata is a mandatory accompaniment to all GIS data transmittals. Data set metadata describe the entire content of a particular collection of data. Usually, the collection, or set, is a single map data file. Data set Examples are an ESRI-format shapefile (which is actually three or more distinct files) or a feature class in an ESRI personal geodatabase. Datum and coordinate system must be documented in data set level metadata. In general one is always well-served by following the FGDC standard, the Content Standard for Digital Geospatial Metadata (CSDGM), which is in release 2.0 as of this writing.

The FGDC standard can appear onerous to the casual data submitter. Whether one uses this standard or not, there are some items that should accompany every data submission. These are shown in the following table. Metadata documents must be sent in plain text formats. Optionally, one might also include the same information in formats created by metadata tools.

Topic	***Description***
Data creator	Company, agency, or other organization who created the data in the dataset.
Date created	Date on which the dataset was created, finalized for conveyance
Associated activity, resource identifiers	A list of identifying numbers associated with this dataset. Typically, this might be an organization project number, an agency investigation number, and a SHPO activity or project number. The purpose is to lead the user to appropriate paper records.
Methods, data processing description	Methods and data processing techniques used to create the data. A brief description will suffice This topic can include the field procedures, equipment, and protocols used for collecting spatial data. For GPS data collection, this could include receiver make and model, PDOP cutoff, etc. Post-field processing, aggregation, digitization, and smoothing may be described in this section. Datum conversions performed upon the data set should be specified.
Responsible party and point of contact in creating organization	The name(s) and contact information of a person in the data-creating organization who is familiar with the data and responsible for its quality.
Coordinate system, units, and datum of data	This topic must cover the coordinate system and datum in which the data are conveyed (not necessarily the coordinate system in which the data was created – this might be covered in the "Methods" section of the metadata document, including conversion from the source coordinate system to the conveyed coordinate system). Different agency offices may have specific requirements for data that will be accepted.

BLM has created a standard that incorporates the information above and describes business characteristics of a dataset. This standard fulfills all of the mandatory requirements above (WO-IM-2003-125 attachment 2: Guidance for Managing BLM Data Standards: How to Adopt, Implement and Maintain Data Standards, pages 17-20).

Entity Data and Metadata

The following sections describe mandatory columns that should be present in each GIS attribute table and must be populated with attribute values for each row in the table. For each column attribute, except entity labels, a table of allowed values and what they denote follows the attribute description. For ease of reference, each column is presented on its own page. To accommodate the shapefile format, column names are limited to 10 characters or less.

Entity Identifier(s)

Column name(s): variable, by state-level convention, should be 10 characters or less
Column format(s): variable, by state-level convention
Description: This attribute is actually data about the phenomenon, consisting of one or more identification values. State-level conventions and standards should be used here. For instance, many states use the Smithsonian trinomial numbering system for archaeological sites that have been formally recorded and thus would require this number to be present if known. The Smithsonian number could be in three columns (state identifier which may be numeric or character, county identifier which is typically 2 characters wide except in California, and resource number) or just a single column as in "26WA1234". Similarly, almost every entity has both a field or temporary or local identifier (a temporary site identifier, an agency-specific project number) and a formal identifier (an assigned AHR number in Alaska, an LA number in New Mexico, a Smithsonian trinomial, a California p-number, a master survey report number, etc.). Appropriate identifiers, formats, and allowed values should be identified at the **state** level and are mandatory if applicable.

Attribute Values: State-specific.
Example Values: Using California as an example, Column "PNUMBER" (9 character string, nulls allowed) example values: "16-000123", "32-987654"; Column TRINOMIAL (13 character string, nulls allowed) example values: "CA-LAS-012345".

Horizontal Positional Accuracy
Column name: HposAcc
Column format: character, 6 wide
Description: This attribute describes the *horizontal* positional accuracy of the GIS entity. Accuracy can be conceptualized as the likelihood that a stated coordinate is the true coordinate of a position. Hence, accuracy is the converse of positional error. The values for this attribute are probable positional error circles – the root mean square (RMS) error of a position.

For a single position, the root mean square error is a clear measure of accuracy probability. Many GPS units and post-processing software return RMS errors for averaged position fixes. However, there are many cases in which RMS error is more difficult to determine. RMS error is an estimate derived from repeated measures; single position fixes must default to having an RMS error at least as large as the usual RMS error for the source of the position fixes.

A practical guideline for determining the value of this attribute is the source of the coordinates ("location") used to create the GIS entity. So, a map coordinate measured from a USGS 1:24,000 quadrangle has an RMS error greater than the paper map itself.

The entity-level attribute is important because combining data with different horizontal positional accuracy can yield unexpected results. For example, a highly accurate GPS-determined position may plot on the "wrong" side of a USGS digital map image registered in coordinate space. This is not due to an error in either position, just the positional error in the map is greater than the error in the GPS position. Adjusting the GPS-position to match the map would degrade the horizontal positional accuracy of the GPS reading (and the HposAcc value should be changed accordingly).
Attribute values: Attribute values are the roughly 90% probability that a stated coordinate lies within a certain distance of the true coordinate.

Value	Example methods used to determine coordinate(s)
<1m	Averaged, differentially corrected high-end resource grade GPS; Survey-grade GPS; Experienced operator using 10" or more precise total station or theodolite and EDM traversing from a known coordinate monument less than 5000m distant
<10m	Single position of high-end resource grade GPS; multi-position averaging of sports-grade GPS without differential correction
<20m	Typical sportsman grade GPS – single position fix; USGS 1:24,000 map (National Map Accuracy Standard is approximately 13m)
<100m	USGS 1:36,000 to USGS 1:125,000 map
UnkLow	Unknown – low confidence in horizontal positional accuracy; likely error is not known, location is only an estimate quite likely to be erroneous
UnkHi	Unknown – high confidence in horizontal positional accuracy; likely error is not known, but coordinates are likely to be correct on a 1:24,000 scale map
UnkUnk	Likely error is not known and no estimate of reliability of horizontal position is possible

Example Values: "<20m", used to indicate horizontal position came from a USGS 1:24,000 scale quadrangle.

Horizontal Positional Source
Column name: HposSrc
Column format: character, 10 wide
Description: This attribute describes the source of the coordinates used to place the GIS entity into coordinate space. The attribute values describe only the most common sources and are not intended to be comprehensive. Horizontal positional source is useful as a means to segregate GIS entities derived from different sources, especially in data derived from plots on paper maps.
Attribute values: Attribute value is determined by the *source* of the horizontal coordinates. A GIS entity may have multiple sources, in which case one should state the predominant source. Multiple source entities that have no dominant source should receive an attribute value of “other”.

Value	Example methods used to determine coordinate(s)
GPS	A GPS unit, of any grade, was used
SurvInst	A total station or a survey instrument (transit, alidade, theodolite, electronic distance meter, stadia rod, or chain/tape), was used
USGS map, scale, e.g., USGS24000 USGS62500 USGS63560 USGS100000 USGS…	Horizontal position coordinates were derived from USGS map at given scale. Note that if one transfers a GPS position to a map, then digitizes from the map, the accuracy is still that of the map, not the GPS.
Aliquot	Derived from an aliquot (cadastral) location. This depends upon the size of the aliquot part relative to the entity coordinate. At best, since an aliquot must be mapped to be converted to coordinates, the horizontal positional accuracy is that of the associated map.
Asserted	Horizontal position is an assertion with no other source information (e.g., a site record). In this case, horizontal positional accuracy will probably be unknown.
Other	Some other source, known but not among choices above.
Unknown	Source is not known.

Example Values: “USGS24000”, the source of the horizontal position information was a line drawn on a USGS 1:24,000 scale quadrangle.

Boundary Precision
Column name: BndPrec
Column Format: character, 6 wide
Description: Boundary Precision is the “fuzziness” or uncertainty of a reported boundary. It applies only to polygonal (having the geometric property of area) GIS entities. “Fuzziness” can be thought of as how sharply a bounding line should be drawn. An inaccurate boundary would be represented as a wide gray line, a very accurate boundary as a thin, darker, line. Current GIS display technology does not do a particularly good job of displaying these uncertainties in data. Note that the concept of boundary precision does not, generally, apply to whether a boundary is real, imagined, or how it was estimated. Rather, Boundary Precision is the reliability one places upon the boundary as a set of coordinates. Some other means is necessary to determine whether one should trust the method by which the boundary was defined on the ground.

Boundary precision can be a complicated estimation if one considers all of the potential error sources and uncertainties that compose a bounding line. For example, if one creates a boundary by joining together high accuracy GPS positions, what is the “fuzziness” of the lines between the points? This will depend upon how closely the position fixes fit the intended boundary. Too few points, or points in the wrong place, and a boundary can be quite inaccurate.

Many GIS polygons are composed of heterogeneous boundary sources, each of which could have its own spatial inaccuracy. In the standard presented here, each GIS polygon is given a single value for boundary precision. A more complex standard would involve associating appropriate accuracy attributes with *each* part of a polygon boundary. Although perhaps desirable, individual boundary segment attributes would be complicated to create and manage. For this reason, they are not incorporated into this standard. A single estimation is requested in this standard.

In general, the predominant technique used to gather or create coordinates of the ***observed*** boundary of an entity determines the boundary precision. Estimated parts of boundaries are not to be included in the estimated accuracy. Because of its inherent complexity, boundary precision will always be a judgment of the cartographer creating the GIS entity.

Attribute values: Attribute values are the estimated, appropriate values for a “gray” line to represent the boundary of a phenomenon, were a GIS to draw the boundary as a zone of probability. The values are intended to be best judgement, realizing that one will probably be combining different error widths in most cases.

Value	Example methods used to determine boundary
<1m	Averaged, differentially corrected high-end resource grade GPS; Survey-grade GPS; Experienced operator using 10” or more precise total station or theodolite and EDM traversing from a known coordinate monument less than 5000m distant
<10m	Single position of high-end resource grade GPS; multi-position averaging of sports-grade GPS without differential correction
<20m	Typical sportsman grade GPS – single position fix; USGS 1:24,000 map (National Map Accuracy Standard is approximately 13m)
<100m	USGS 1:36,000 to USGS 1:125,000 map
UnkLow	Unknown – low confidence in horizontal positional accuracy; likely error is not known, location is only an estimate quite likely to be erroneous
UnkHi	Unknown – high confidence in horizontal positional accuracy; likely error is not known, but coordinates are likely to be correct on a 1:24,000 scale map
UnkUnk	Likely error is not known and no estimate of the accuracy of horizontal position is possible

Example values: “<20m”, indicating that the horizontal imprecision of the boundary is 20m or less.

Boundary Observation Completeness

Column name: BndComp

Column format: character, 8 wide

Description: The Boundary Observation Completeness Attribute describes whether the GIS entity boundary is the entirety of the boundary of the entity being mapped or only part of the entity boundary. The attribute is particularly useful in situations where only part of a phenomenon (e.g., a resource, an investigation) is mapped in the field. The attribute flags the observational completeness of the phenomenon boundary representation, not the logical completeness of the boundary. A boundary is logically complete simply by closure (for a polygonal entity); observational completeness means that the logical boundary matches the actual boundary.

Note that in the case of a linear or point phenomenon, an observed boundary may take the form of a line (perhaps the centerline of the phenomenon) or a point (perhaps the central point in the phenomenon).

In practical terms, someone using the spatial data is given a means to determine whether the data are complete for a given phenomenon or whether the boundary shown is closed merely by convention. This determination is often very important for “linear” entities, such as roads, trails, or ditches. Segments of these phenomena may be recorded in their entirety (see below under the Segment attribute), even though the entire road, or ditch, or pipeline is not completely observed. Yet, each segment recording is a complete observation. No part of the reported segment boundary is an inference.

Attribute values: Attribute values for Boundary Observation Completeness signal to the user whether a boundary was completely observed or not.

Value	Boundary observation completeness
Complete	Entire phenomenal boundary was observed (mapped) completely. Note that this could mean the centerline of a phenomenon, or a centerpoint. By convention, a distinctly identified segment of a phenomenon can have a value of "Complete" if it has been mapped entirely.
Partial	Only part of the phenomenon was mapped.
None	The phenomenon boundary was not observed or mapped at all. Boundaries created by buffering using a convention (e.g., "sites less than 30m in extent shall be mapped as point and buffered to be a polygon 30m in diameter") would have a value of "None" for Boundary Observation Completeness.
Unknown	The observational completeness of the phenomenon spatial data is not known.

Example values: "Partial" would be used in the situation described below, because part of the resource lacks boundary definition.

Consider an archaeological site recorded within a highway right of way. The crew recording the site is not allowed to leave the right of way, although the site runs outside of the right of way. So, the crew maps the boundary of the site as they observe it right up to the edge of the right of way. In the GIS, the polygonal shape representing this archaeological site is "squared off" at the right of way edge – the GIS entity is logically complete. In other words, there is a boundary represented in the GIS, but the entire boundary was not observed, so it does not represent the boundary of the entire phenomenon.

Segment Status

Column name: Segment

Column format: character, 1 wide

Description: The segment variable describes whether the spatial entity represents a definable segment, lobe, or part of an entity, rather than the entire entity. It is necessary because some phenomena are so extensive spatially, forbidding of access for mapping and observation, or otherwise unobservable. A spatial data user may be presented with a spatial entity of high positional and boundary accuracy, complete boundedness, and yet be shown only part of the entire phenomenon. This attribute flags such a condition for the spatial data user. If true (the entity is a segment), then there is more of the same phenomenon, perhaps present in data as different spatial entities. If false, then the entity represents the entire phenomenon.

The State of Hawaii is an excellent example of the use of Segment. In a GIS, the state is represented as several distinct polygons. Each is a part (a segment) of the state but is not the entire state. Thus, Segment = TRUE for each island.

Linear phenomena are particularly amenable to segmentation in spatial datasets. Simple examples abound: the portion of a highway that lies within a particular county. A railroad construction shoo-fly.

A lobe of an entity is also a segment. Above, the example of an archaeological site in a highway right of way was given (see Boundary Observational Completeness). The portion of the site within the right of way is a lobe, or segment, of the entire site. So, not only was the boundary incompletely observed, but the spatial data represents a segment of the entire phenomenon.

There is no necessary relationship between Boundary Observation Completeness and Segment. A boundary may be complete, but the entity is only a segment (e.g., a single Hawaiian island). A boundary may be incomplete and the entity is only a segment (the highway right of way site example). A boundary may be complete and the entity is not a segment (the entity represents the entire phenomenon). A boundary may be incomplete and the entity is not a segment (for example, a partly observed archaeological site bound).

Attribute values: Attribute values for Segment are straightforward.

Value	Segment
T-TRUE	The entity is a segment of the entire phenomenon.
F-FALSE	The entity is the entire phenomenon.
U-UNKNOWN	The relationship of the entity to the phenomenon is unknown.

Example values: A portion of an historic railroad grade is present on both sides of an interstate highway, having obviously been destroyed in the interstate roadbed itself. A field crew would record each piece of railroad grade with values of SEGMENT = TRUE, because neither piece is the entire railroad grade.

Maximum Entity Width (meters, applies only to point and linear entities)
Column name: EntWidM
Column format: integer
Description: In many cases, phenomena are recorded in GIS as points and lines, even though they are two-dimensional (i.e., they have area). The Maximum Entity Width column gives a single metric value representing the width of the entity. If one were to create a spatial boundary around the entity, then one would use half the Maximum Entity Width as a buffer distance.

When a phenomenon is presented in GIS as a polygon, it will not have an entity width (the value should be zero).
Attributes: The values in this column are the actual width or diameter of the phenomenon's spatial extent.

Value	***Description***
0	Entity should not be buffered to create a polygon (e.g., it already is a polygon)
1 to any value	Width or diameter of resulting polygon if entity is buffered to create phenomenon of appropriate size. Buffer distance in most GIS software would typically be half of this value.

Example values: EntWidM = 20 -- A small lithic scatter is recorded as a single x, y position. The small site extends over an ellipse of roughly 20m by 10m. The maximum dimension of the site is 20m, so a value of 20 is used. When a GIS user wants to create a circle containing the reported site, the x, y location is the center point of a circle with a radius of (EntWidM / 2). The small site will lie inside of this circle.

Summary

This standard has only five major elements. These are shown in the table below. Each element has a role in the effective use of spatial data.

Element	**Requirement**	**Rationale**
File format – shapefile	Recommended	Easily created by wide variety of software, non-topological
Datum and Coordinate System – NAD83, UTM or latitude-longitude	Mandatory	Variation is allowed but must follow defined practice at state or regional level
Dataset spatial metadata	Mandatory	Required as part of agency business process and as best practice
Entity spatial metadata and identifier(s)	Mandatory	Required to determine utility of information for business purposes

To facilitate implementation of these standards, examples of entity spatial metadata have been prepared in shapefile and ESRI ArcGIS personal geodatabase format (release 9.1). These are available at:

ftp://www.gnomon.com/BLM_GIS_Standard_Examples

The example files contain the standard columns discussed above, as well as columns for resource centroids (a mandatory observation for most resources per BLM instructional bulletins), some example identifier columns, and some information on who created the record and quality control activities. The personal geodatabase contains attribute domains for the spatial metadata columns that match those shown above.

Appendix A. GPS, Cartography, GIS, and Metadata

UNITED STATES DEPARTMENT OF THE INTERIOR
BUREAU OF LAND MANAGEMENT
WASHINGTON, D.C. 20240

October 20, 2003

In Reply Refer To:
8100 (240)P

EMS TRANSMISSION 10/22/2003
Instruction Memorandum No. 2004-020
Expires: 09/30/2005
To: All Field Officials
From: Assistant Director, Renewable Resources & Planning
Subject: Guidance for Recording Cultural and Paleontological Resource Locations for the Bureau of Land Management (BLM) using Global Positioning System (GPS) Technology
DD: 04/01/2004

Program Areas: Cultural and Paleontological Resources

Purpose: The purpose of this guidance is to provide a minimum set of requirements for recording cultural and paleontological resource locations for the BLM using GPS technology. The GPS has become a major tool for Geographic Information System (GIS) and traditional mapping applications. The use of GPS technology to record all site locations for the BLM shall be required within six months from issuance of this Instruction Memorandum.

The main objective of this guidance is to improve the overall reliability of site location information recorded by field archaeologists, paleontologists, and other specialists working within the BLM or working on lands administered by the BLM, including contractors; and support the standardization and expansion of GIS applications for cultural and paleontological resource management.

Policy and Action: This guidance is intended to produce overall cultural and paleontological resource location data with a mean error of +/-12.5 meters or less, at a 95 percent confidence level. The mean error requirement is consistent with the National Map Accuracy Standard for 1:24,000 scale quadrangles and Federal Geographic Data Committee (FGDC) reporting requirements. This accuracy can be achieved with a variety of contemporary GPS equipment. Appropriate equipment is defined as GPS technology that meets the accuracy standard.

Cultural resources shall be located by reporting a minimum of one GPS-observed coordinate taken in the approximate estimated visible center (centroid) of the resource. The centroid need not be perfectly central to a site, but it must lie in the site's approximate center for map-plotting purposes. Multiple coordinates shall be used to
define the approximate centerline of a linear resource (e.g. trail), if field judgment suggests that a single centroid is insufficient to record its location. More points, lines or polygons may be taken for other mapping purposes, including recording project area boundaries, site datums or markers,

or internal attributes. Applicability of this standard for recording isolated finds shall be a state-level decision.

Paleontological resources shall be located according to the guidelines set forth in the BLM Handbook H-8270-1, General Procedural Guidance For Paleontological Resource Management, Ch. II A(4) and Ch. IV P(1) and expressed in Universal Transverse Mercator (UTM) North American Datum 1983 (NAD83) coordinates. Points may be used to identify discreet sites or isolates; lines or polygons may be used to delineate site or project boundaries.

Archaeological resource locations shall be reported in an appropriate, identified, coordinate system. The BLM's standard for coordinates is Universal Transverse Mercator, North American Datum 1983 (NAD83); whenever possible, coordinates should be reported using the NAD83 values. However, standards may differ between States and in collaboration with State historic preservation offices; consequently, all reported coordinates must clearly identify the coordinate system used.

In situations where GPS observations are not practical or possible due to geography, vegetation, satellite availability, or the presence of hazardous materials, the recorder should locate the resource using GPS offset equipment and capabilities, map coordinates, or a combination of GPS and other techniques. Such non-GPS methods must be described in the site or project area record.

The GPS observations will be reported on the appropriate part of a resource recording form, in the narrative description of the resource, or both, and include the following information:

- The UTM coordinates with the UTM zone should be reported. For all coordinates, the datum reference must be reported.
- The coordinate system for observations should be recorded in an obvious way (e.g. "UTM Zone 10 NAD83 centroid coordinate: N4986000 E302000 meters")
- If the error terms for a given coordinate are known, then the probable error must also be recorded in narrative (e.g., "GPS observations were differentially processed to an average error of less than 5m root mean standard deviation [RMS]").
- Receiver type, correction status, length of observation and number of observation points, position dilution of precision (PDOP), and horizontal error estimates must

be recorded with the location whenever GPS equipment and software provides such information.

- Discrepancies between GPS locations and USGS quadrangle locations should be noted on the site record. Because GPS locations are mathematically precise coordinates, a point plotted from GPS may appear to be in an incorrect location on a USGS quadrangle.

This is a minimum standard and should not be used to lessen any applicable State, agency or Federal standard or reduce site location accuracy from conventional mapping methods. There will be situations where more accurate location information is desirable, or required. For instance, State Offices may apply more stringent standards for intra-site mapping, excavation unit and datum locations. In all instances, the most accurate and capable equipment available shall be used to meet the needs of the types of data that are being recorded, even if it exceeds the accuracy suggested in this guidance. Appropriate GPS experts within Washington Office, National Centers, State and Field Offices should be consulted as needed.

Timeframe: This minimum requirement for recording cultural and paleontological resource locations for BLM using GPS technology is in effect on April 1, 2004.

Contact: Please contact either Marilyn Nickels, at (202) 452-0331, or Linda Clark, at (208) 756-5460 with any questions.

Signed by:
James G. Kenna
Acting Assistant Director
Renewable Resources and Planning

Authenticated by:
Barbara J. Brown
Policy & Records Group, WO-560

The GPS Revolution

GPS is revolutionizing how maps get made. Even low-cost receivers can calculate geographic coordinates more precisely than one can determine them from a 1:24,000 USGS topographic map. GPS information can usually be converted into a GIS file format compatible with BLM and SHPO spatial databases. Cultural resource information systems (managed by BLM or by SHPOs) can benefit greatly from streamlined transmittal of *appropriate* digital data.

GPS field recordings, even those made with sophisticated data collectors, cannot usually be imported into GIS data systems without some manipulation. The difficulty is not file formats, but in the nature of GPS data itself. GPS is a mapping tool, like a theodolite, an electronic distance meter, or a stadia rod. These tools require skillful manipulation of their results to create a coherent site map. So does a GPS. GPS does make collecting the information needed to make effective maps much easier.

Managers of cultural resource information systems (in SHPOs or in agency offices) have little interest in receiving all of the GPS files from a fieldwork episode, just as they had no interest in receiving the instrument books from people mapping with alidades and plane tables. Cartographic representation was and is an interpretive activity. Managers of cultural resource information systems need to receive cartographic representations of cultural resources phenomena in a format appropriate for error checking and rapid inclusion in their information systems.

Coordinate systems are a fundamental aspect of spatial data. GPS units collect coordinates in latitude and longitude, or spherical coordinates, by default. Except for spherical coordinates systems, all coordinate systems in some way fit the curving surface of the earth to rectilinear, planar , axes. Every coordinate system is based upon a datum point, which in turn is linked to a particular models of the earth's shape. The North American Datum of 1927 (NAD27) is based on a simple elliptical model of the earth first formulated in 1866. Almost all published, paper, maps in the United States are based on NAD27. A more complex model of the earth resulted in the creation of the North American Datum of 1983 (NAD83). A large body of digital mapping data has been collected in NAD83, and GPS units use this datum (or technically its global equivalent) by default to report coordinates. The BLM and most federal agencies are using NAD83 as the digital map datum, though NAD2004 will soon be available. BLM's standard is currently NAD83, and a best practice is that this be the preferred datum in transmitted data sets. Most GIS programs allow for reprojection or direct use of geographic data as if it was in a planar coordinate system. Thus, the most universal format is geographic coordinates.

Cartography and GPS

Most agency offices currently map site locations and boundaries to a 1:24,000 map base in the lower 48 states and 1:63,560 map base in Alaska. While there are standards for every formal map series published by the U.S. Geological Survey, the general notion in agency map files is that the cultural resources depicted are "pretty good" representations of site location and boundaries. This is inherent in the map scale. For example, a 0.5mm pencil line on a 1:24,000 map is 12m wide. A site represented as a visible 2mm dot would be 48m in diameter and cover an area of about half an acre (actually 1,806 square meters or 19,516 square feet).

GPS brings much higher accuracy to the recording process. Even uncorrected GPS from "sports model" GPS receivers will plot accurately at 1:24,000 scale. At larger scales (e.g., 1:6,000), inconsistencies will start becoming evident in uncorrected data.

Fieldwork gathers GPS and other spatial information in whatever way is most effective. Sites may be mapped using a variety of tools at the same time resulting in raw spatial information. In turn, the field mapping information is cleaned and made more regular to produce site "sketch" maps at scales of 1:100 to as small as 1:24,000 and site "location" maps at scales of 1:24,000 for inclusion in the site record packet.

For paper cartographic purposes, the draftsman can synthesize multiple sources of spatial information into a coherent single map. In digital cartography one is always tempted to retain the original digital data to the fullest possible extent, yielding a map that is complex and possibly difficult to interpret. For example, a site boundary could be created from GPS lines, GPS points, interpolations between points and lines, topographic lines from a USGS map, and a fenceline traced from a digital aerial photography. If one asked, “what is the spatial accuracy of the boundary?” there would be no simple answer.

An Introduction to Metadata

The standards proposed in this document are use-oriented. That is, they aim toward the use of individual pieces of information rather than documentation of its genesis. The user or evaluator of spatial information that follows metadata standards should be able to answer a question posed about the spatial accuracy of a boundary like “how spatially accurate is this site boundary?” . This is because the standard presented in this document creates *metadata* (literally, data about some item or body of data) for each spatial entity, as well as for a GIS dataset as a whole.

Data are primary pieces of information about the phenomenon itself. For example, a site identifier is a primary piece of information about a site. Metadata are simply additional pieces of information that allow the user of data to better understand the primary information (i.e., data) that they are using. For example, knowing that the site data (primary attributes) were entered from field notes rather than a formal site record might lead you to consider it less reliable for some purposes. Metadata are particularly useful when it comes to spatial information. For instance, estimating the likely accuracy (or inaccuracy) of a resource’s spatial location is very useful when using a map that shows where the resource is (or is said to be).

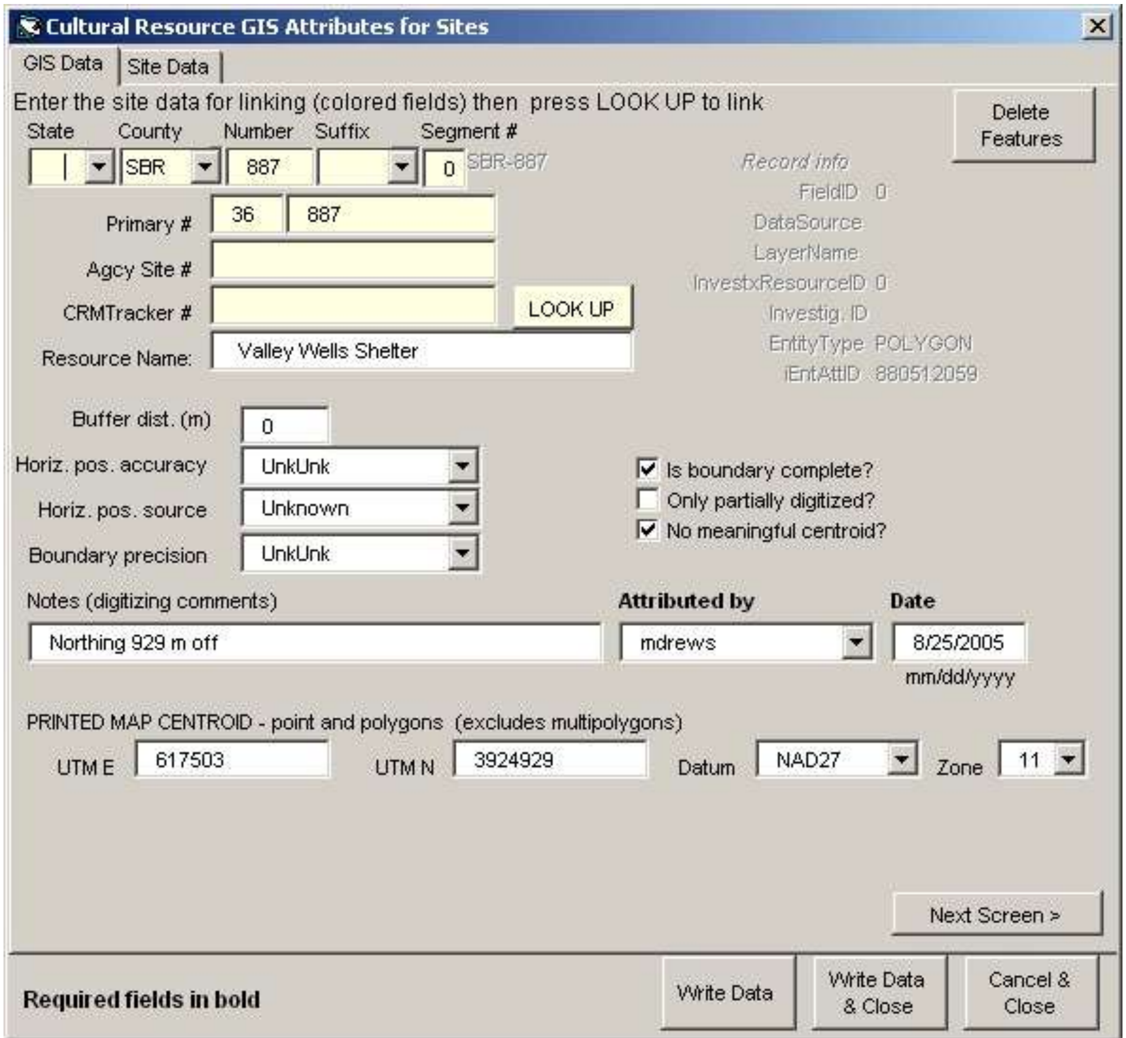
However, one must recognize that the distinction is artificial in some important ways. For example, is knowing the date a site record was filled out data or is it metadata? It is fairly obviously the former, but isn’t it also the latter because terms, field methods, and standards for recording change over time?

There are many approaches to metadata. One is to describe the full heritage of each data item. Another approach is to describe entire sets of data. The latter approach is used in current Federal Geographic Data Committee standards for geospatial metadata. This data set standard is still part of this proposed standard.

A second approach is to describe individual entities within a data set. This approach, which is also called feature-level metadata, is appropriate when data are heterogeneous in source, origin, and variable in data quality. Because each entity in a cultural resources information system may have come from different fieldwork episodes, mapping techniques, and digitization methods, this standard focuses on the characteristics of individual GIS entities, the proposed standard incorporates entity-specific metadata as well as data set metadata.

Appendix B. Example Interface for Metadata Entry and Editing

The figure below displays the entry screen for the current version of the BLM California Arch Edit tool. This GIS toolbar runs in ArcView 9.x. When a cultural resource feature is selected with the tool, the screen below is presented (for resources – a similar screen is used for investigation entities).



The screen displays many of the features discussed above. In the upper left of the screen, *identifiers* are entered. In the middle of the screen, the *buffer distance* is requested, as are *horizontal positional accuracy, horizontal position source*, and *boundary precision.* In the middle of the screen, *boundary completeness* and *segment status* (phrased as partial digitization) are gathered as check box entries. The rest of the screen shows some other useful values that are not discussed as part of this standard. In practice, the use enters relatively little information because most values are defaults.

The BLM California application uses a personal geodatabase as its data storage. The geodatabase contains all of the lookup values for the metadata and the data itself. This enforces standardization of terms, even if someone edits the attributes without using the form shown above.