Guidelines on Information Deliverables for Research Projects in Grand Canyon National Park

INTRODUCTION
Science is playing an increasing role in guiding National Park Service (NPS) management activities. The NPS is charged with protecting and maintaining data and associated information that have been collected through scientific research so they can be used to assist new investigations and/or inform current and future management decisions. This document provides guidance to individuals planning to conduct research at Grand Canyon National Park (GRCA) about NPS data policies and standards.

GENERAL REQUIREMENTS
At the completion of a study, researchers are requested to submit to the park two paper (one unbound) and one electronic copy of all final reports, publications, and theses/dissertations produced as a result of the permitted project. Copies of all associated data (including field notes, maps, slides, photographs, charts/graphs, tabular and GIS data with associated metadata) are also requested to be submitted to the park. These materials will be permanently archived in the park's museum collection.
NOTE: If your project is federally funded, you are required to submit these materials and may have more specific formatting requirements.

To facilitate project administration and effective communication, a park staff member may be assigned as a liaison to each permitted project. If deemed appropriate, prior to the start of a study in GRCA, the park liaison will contact the researcher to review the project activities and types of information that will be collected, coordinate any logistical issues, and discuss the products that will be delivered to the park.

Grand Canyon's Data Manager (Santiago Garcia; 928-638-7450 or Santiago_Garcia@nps.gov) and GIS Coordinator (Mark Nebel; 928-638-7451 or Mark_Nebel@nps.gov) can provide further guidance on data formats, metadata requirements, and product delivery schedules.

Investigators who plan to collect specimens that will be archived and/or displayed in a museum collection, must also contact the park Museum Curator (Kim Besom; 928-638-7769; Kim_Besom@nps.gov) **PRIOR** to initiating any collecting to discuss collecting responsibilities (such as specimen preparation and cataloging) and ultimate disposition of the specimens.

DELIVERY
At the conclusion of the project, researchers will provide the park liaison with project data and information previously agreed upon, which may include:

- **Written reports, publications, theses/dissertations, and other manuscripts** (two copies of each, one of which is unbound; we also request one electronic copy – preferably .pdf or MS Word format)
- **Raw data** (including field data sheets, field notes, lab reports, drawings, maps, literature reviews/bibliographies, databases/spreadsheets, etc.)
- **Analyzed data** (including reports, charts, and graphs)
- **GIS data**
- **Tabular and GIS metadata**
- **Photos and negatives; video and audio tapes** created during research activities
- **Specimens** and species lists

Note: Please submit all digital data or applications to Grand Canyon National Park on media that are compatible with our current hardware (CD-ROM, DVD, posted for ftp retrieval, or attached to email).

SPECIFICATIONS
TABULAR (NON-SPATIAL) DATA
The NPS-preferred storage of raw field data is in relational MS Access 2000 database. The park's Data Manager can provide further suggestions on specific aspects of database design and structure that can make your data more useful to park managers and future researchers. Additionally, the NPS Inventory and Monitoring Program has developed a

database template (found at http://science.nature.nps.gov/im/datamgmt/applications/template/index.cfm) that can be used for your study data and will ensure that they conform to NPS standards.

Although Access is the NPS database standard, data stored in commonly available software such as Excel, dbf, delimited text, etc. are also suitable for submission. Electronic data are requested to be submitted to the park in both digital (CD preferred) and hardcopy (paper) form.

Tabular data should include metadata conforming to the specifications found below. Metadata requirements are similar to the Federal Geographic Data Committee metadata standard for geospatial data. Complete metadata will enhance the usability and longevity of study data.

METADATA
Metadata will be created in Microsoft Word. The Word document will consist of the elements listed below along with a narrative for each element. Metadata will consist of information describing the following elements:

- Data set name
- Purpose of data set
- How data set content was created
- Person(s) responsible for data set
- Time period of data set content
- Time period reference
- Keywords describing the data set
- Keywords describing the spatial location of the data set
- Data quality assessment
- Access constraints
- Data set format
- Data set retrieval information
- Contact for further information

Purpose of data set
Describe why the data were/are collected. The intended use of the data set will help the reader assess how it may be utilized.

How data set content was created
Explain how the data set content was created. If data were collected in the field describe the method(s) or protocols used.

Person(s) responsible for data set
Identify the individual(s) or group that created the data set or collected the data. It may be more appropriate to name the SRM Program or project name for projects that involve many people.

Time period of data set content
Provide date(s) the data set was created or dates during which data were collected. Provide specific dates or ranges (e.g. 08/24/2012 – 11/15/2012) for data sets with finite dates or specify a range of years (e.g. 2000 – 2012) for data sets spanning multiple years.

Time period reference
The basis for which the time period of content is determined. Valid values include: ground condition, publication date.

Keywords describing the data set
Assign multiple keywords to a data set. Keywords will provide a mechanism to search data sets. A standard list of keywords will be available and maintained in the Data Registry. In the meantime, use descriptive keywords.

Guidelines on Information Deliverables for Research Projects in Grand Canyon National Park

Keywords describing the spatial location of the data set
Assign at least one of these keywords. These keywords provide an indication of where the data are collected:
Above Rim, Inner Canyon, River Corridor, Developed Area, Outside Park

Data quality assessment
Data quality will be addressed in terms of these criteria – complete, correct and reliable. For each criterion provide a qualitative assessment and then an overall data quality assessment. The overall assessment will include a rating – high, medium or low.

Access constraints
Describe access and distribution restrictions. Identify groups or individual staff that are permitted access and what type of access (edit or read only). Describe constraints to distributing data both within and outside of NPS.

Data set format
Identify the format of the data set. A partial list of valid values includes:

Access database	Comma-delimited text file	Excel spreadsheet
GIF digital photo collection	GIS feature class	GIS mosaic dataset
Hard-copy document collection	Hardcopy photo collection	JPEG digital photo collection
MP3 audio file collection	PDF collection	SQL Server database
TIFF scanned documents	WAV audio file collection	Word document collection

The list of data formats will grow as new formats are used.

Data set retrieval information
Provide technical details on how the data set is retrieved or accessed. For example, provide the URL for a web application designed to access the data set. *This information will not be available for restricted data sets.*

Contact for further information
Provide contact information for personnel that can provide further details. Include name, phone number and email address.

Data Dictionary
A data dictionary provides detailed information on the fields found in a data set. Data dictionaries will be created in Microsoft Word. The Word document will consist of the elements listed below for each field. Fields which are self-descriptive are not included. For each field the following information is provided:

- Field name
- Field description
- Data type
- Value domain

Field name
The field name as spelled in the data set.

Field description
Describe the field and the information captured by the field.

Data type
Identify the data type. Fields found in databases have an explicit data type (e.g. integer or varchar(45)). Fields in spreadsheets do not require data type definitions. In those cases chose one of the following – text, integer, decimal or date.

Guidelines on Information Deliverables for Research Projects in Grand Canyon National Park
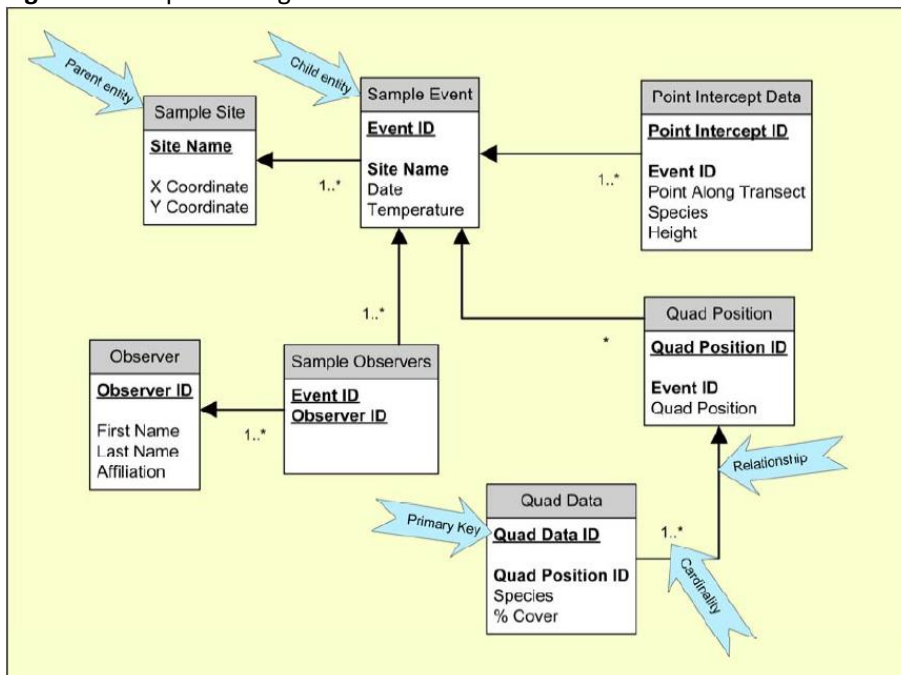
Value domain
Some fields are limited to a domain or set of values. Provide the list of valid values. In cases where the domain is a large list provide a reference to a document actually listing those values.

Entity Relationship Diagram (ERD)
ERDs will be created for data sets stored in databases (Microsoft Access or SQL Server). ERDs utilized for documentation will reflect the logical data model. ERDs will be created in either Microsoft Access or SQL Server, depending on the database. Both software include tools to create ERDs. The diagram will be created and maintained in the database software and exported as a graphic (GIF) for inclusion as documentation. Figure 1 shows a sample ERD that would be suitable for documentation. The ERD depicts all data tables, primary and foreign keys and relationships between tables.

Figure 1 Example of a logical data model

GIS (SPATIAL) DATA
You must submit final GIS data layer(s) you developed as a result of your research or investigation at Grand Canyon National Park. Publicly-available GIS layers used in the study, e.g. NHD hydrology, GCDB land survey, USGS topographic maps, etc., do not need to be submitted unless modifications relevant to your study have been made.

Geospatial Data Requirements
 All geospatial data or applications delivered to Grand Canyon National Park must be compatible with the park's current software (ESRI ArcGIS version 10.1), and must be projected in the Park standard coordinate system.

Vector geospatial data
Feature classes in the ESRI File Geodatabase format are preferred. Intuitive field names and attribute values are greatly preferred over coded values. Personal geodatabase feature classes are acceptable if file size does not exceed 300MB. ESRI shapefiles are acceptable only if they are properly projected and only if intuitive field names and attributes can be used without truncation. To ensure integrity of geospatial files during transfer, they should be archived with a Zip utility (.zip).

Raster geospatial data
The preferred format is the ESRI File Geodatabase raster. Other geo-referenced ArcGIS 10.x compatible raster formats are acceptable, provided that projection files to the Park standard coordinate system are included.

Guidelines on Information Deliverables for Research Projects in Grand Canyon National Park

<u>Coordinate System</u>
All submitted GIS files <u>must</u> be projected into Universal Transverse Mercator coordinates, Zone 12 North, North American Datum of 1983 (UTM Zone 12N, NAD83).

<u>Metadata</u>
All GIS data layers created by research funded by federal agencies <u>must</u> be documented with metadata conforming to FGDC (Federal Geographic Data Committee) standards that, at a minimum, include: Dataset description, summary, credits, purpose, sources, keywords, use limitations/constraints or known shortcomings, appropriate scale range, relevant dates and temporal extents, points of contact, data quality, lineage (process steps), field definitions, attribute definitions, and attribute value domains. Metadata directly incorporated into each data layer (ArcGIS Metadata format) via ArcCatalog is preferred, but stand-alone metadata files in eXtensible Markup Language (.xml) format are acceptable.

PHOTOGRAPHS
Our preference is for non-digital photographs and images (slides or prints, color or monochrome). If digital photographs are provided to the park, they should be in .jpg formats, with a minimum resolution of 300 dpi. Digital photos should be submitted on CD or DVD, along with hard copies printed on photo paper or acid-free paper. All photos should have accompanying metadata. Metadata specifications are provided above.