Cave Resources Data Synthesis and Cave Resource Assessment in Grand Canyon National Park

By

Vicki Wallis and Timothy Gibbs

Submitted to:

NPS ATR/Key Official Jennifer Dierker and
NPS Project Manager/Subject Matter Expert Steve Rice
Grand Canyon National Park
1824 S. Thompson St., Flagstaff, AZ 86001

Submitted by:

Ted Neff, Ph.D.
Archaeology Projects Principal Investigator
Museum of Northern Arizona
3101 North Fort Valley Road
Flagstaff, Arizona 86001

Final: January, 2013

Task Agreement and Requisition # R8213100093

MNA – 43

MNA Project Number 100-175-363

Introduction

Grand Canyon National Park (GRCA) is not considered to be a "cave park" such as Mammoth, Carlsbad, or Wind/Jewel, but nevertheless contains substantial cave resources of both natural and cultural significance. The vast and rugged park contains an undetermined number of caves, estimated by some to potentially exceed 3,000. Currently, Grand Canyon has information at varying levels of detail on perhaps 10% of the total. Records of these caves consist mainly of historic paper copies in a series of files.

As a part of managing its cave/karst resources, GRCA worked with the Museum of Northern Arizona (MNA) to: (1) Scan historic cave records into digital format; and (2) Input historic and current cave data into the SQL Server Cave Database. As originally conceived, the project also specified that MNA would undertake a third task consisting of the analysis of database records and subsequent preparation of management recommendations to NPS. The analysis of cave data would include specific tasks such as: Production of maps showing location and distribution of known caves in GRCA; Production of maps showing caves containing archaeological and/or paleontological resources; Production of maps showing caves with non-technical (no rope) access; and Production of maps showing locations with exceptionally fragile cave formations. Due to delays with the development and implementation of the SQL Server Cave Database this third task could not be accomplished. However, as the project progressed it was realized that a necessary component of the analysis of GRCA cave data was the preparation of a geodatabase. MNA worked with GRCA to develop a cave data geodatabase and this work constituted the third project task, along with the scanning of historic cave records and data entry into the SQL Server Cave Database.

Summary of Activities & Methods: Scanning, Data Entry, and Geodatabase Development

Processing of the historic data related to cave resources in Grand Canyon National Park involved three primary activities: (1) The scanning of historic records into digital format; (2) The input of data derived from these records into the SQL Server Cave Database (called the "Troglobytes" database); and (3) The development and initial implementation of a cave geodatabase.

Scanning

The first step in the scanning of the legacy data was a cursory organization of the material. This was accomplished using the general file structure existing in the paper files. Items related to specific caves were consolidated and assigned a Cave ID using guidelines set forth in the original Scope of Work to label folders. When deviations from naming protocol were required, these were agreed upon with the GRCA Subject Matter Expert (SME), Steven Rice. Other records were sorted using existing file titles and divisions. Attempts were made to place unfiled documents in the proper folders. When questions arose, the SME was consulted. Folders and labels were added when necessary, but records were retained in the original boxes. A digital file structure was created, mirroring organization of the physical files.

Once organization was complete, items requiring special handling were removed from folders to be processed accordingly. Oversized paper maps were transferred to the NPS, as a scanner capable of processing larger format articles was required. These were retained by the

NPS once scanning was completed. All photographic materials were segregated and sorted by cave and media type: slides, negatives, color and black-and-white photographic prints.

Digitization of photographic records was accomplished using an Epson Photo scanner and software provided by the Museum of Northern Arizona Archives. All media were handled using practices approved and demonstrated by the MNA archivist. The goal was to produce the highest quality scan possible, while maintaining the current condition and integrity of the historic records.

Naming of digital files resulting from the photo scanning process, and creation of a supporting file structure, was also accomplished with the input of the archivist. This involved assigning a file name that included the Cave ID and any identifying data that was available with the item. Such data might include a date, subject, or location for the photograph, but often did not. In such cases, no assumptions as to the content were made. Only existing information was applied to the file in the form of the filename. This was sometimes merely the number of a slide or negative. Additional identification of content in image files was not attempted . Descriptive annotation will need to be provided by the SME or other qualified expert, if desired. Photographic images that already existed in digital form were integrated into the digital file structure, but names were retained as received. When completed, digital images were moved to the appropriate individual Cave folder. Original photographic media were returned to the physical files.

Paper files related to the caves of Grand Canyon National Park include a variety of items ranging widely in age, size and condition. As with the photographic media, documents were handled with care, using approved archival techniques. Within each individual folder like items were grouped together into categories such as correspondence, trip reports, articles, etc. Duplicates were excluded. All paper clips and staples were removed. In agreement with archivists at the South Rim, these were not replaced. Documents were then separated as to size: letter, legal, oversized, and so on, so as to be scanned most efficiently. Searchable .pdf files were created. Scanned pages were grouped into files of similar items. For example, a file named CAVDOM_Correspondence_Letters.pdf, contains all letters of correspondence found in the Cave of the Domes physical file. Document files for this project range in size from one to two hundred pages.

Approximately the first two-thirds of the scanning was accomplished at the NPS Grand Canyon National Park Resource Management Offices in Flagstaff utilizing the NPS HP Scanjet N9120 Scanner, HP Software and Adobe Creator. This work was scheduled in accordance with the availability of the workstation. Both automatic feed and flatbed options were employed based on the nature of the document to be scanned. During and after a major remodeling and move of the NPS Offices, the HP scanner was not available, so the remainder of the document scanning was accomplished with a Xerox Documate 3640 Scanner and Nuance Scansoft Paperport software. This equipment was provided by the Museum of Arizona to complete the project.

Post-scan processing to guarantee document quality was performed when required. As previously noted, the condition of original paper records is highly variable. Many are extremely old and fragile, others are copies of copies of copies with reduced legibility. An attempt was made to enhance clarity when necessary, but this was not always possible given the status of the original. Random test searches were conducted to be certain that text was

recognized when a search within the document was initiated. All paper records presented were scanned into digital format.

Finally, scanned documents were organized within the digital file structure. The data files are arranged in three major categories: "Cave Files," "Areas With Caves," and "Topic Files." Within the Cave Files are folders for each of 365 individual caves, named according to NPS guidelines. Some empty files exist, representing caves for which paper folders were present, but which contained no documents or images. Folder contents are further divided into subfolders for documents, and the various types of photographic media such as prints, slides, etc. Folders representing Areas with Caves were given names along the same general guidelines as those for individual caves for consistency. There was no requirement to use this coding for any reason. Subfolders for each Area With Caves are the same as those for individual caves. Topic Files include all general administrative and informational paperwork pertaining to Cave Management in Grand Canyon National Park. The great majority of files in this category are documents. In addition to correspondence of various kinds, there are publications of caving organizations, scientific papers and research proposals. Divisions in this category follow those of the physical files.

Scanning activities resulted in the following totals:

File Type	**Approximate Number**
Document	**700**
Photographic Image: Photo Negative	60
Photo Print	360
Digital Photo	140
Slide	760
Photo Proof	1
Total Images	**1,321**
Total Digital Records	**2,021**

All legacy data files were transferred to Grand Canyon National Park on August 8, 2012. All digital files were transferred on October 2, 2012, bringing this portion of the project to completion.

Data Entry

Information contained in the scanned legacy documents was used to populate the Troglobytes Database. Performance of this portion of the work was delayed until the NPS had received a final version of the database from its designers. Once approval was received from the GRCA SME, training was provided and work scheduled when an NPS workstation with the required hardware and software was available for use. Inventory Assessments were created for all caves that had either recent Rapid Inventory Assessment Data Sheets available, or enough pertinent information in legacy data to complete an entry. Data was not available for all 365 caves with folders in the resource files.

As problems with functioning of the database, issues with quality of data, and questions about consistency of entry were encountered, they were discussed with the SME. Agreement was reached as to the handling of each issue. In particular, when questionable data was noted on a data sheet, it was decided that this would be entered as it existed, to be replaced by more current information when obtained. Some suggestions were provided to NPS database specialist for alterations and corrections to the database.

At conclusion of the project and transfer of products to GRCA, we were informed that the Troglobytes database was not operational to such a level that all of the products detailed in the original scope of work could be produced. However, all searchable .pdf and high resolution .jpg files, organized as agreed upon, were delivered, ready to be utilized in whatever manner is desired in the future.

Geodatabase Development and Initial Implementation

The analysis of cave data to produce maps showing location and distribution of known caves in GRCA, maps showing caves containing archaeological and/or paleontological resources, maps showing caves with non-technical (no rope) access, and maps showing locations with exceptionally fragile cave formations is greatly aided by a geodatabase. Therefore, MNA worked with GRCA on the development and population of a cave geodatabase. Geodatabase development included georeferencing and digitization of all GRCA caves and karst data.

A tracking matrix was developed in Microsoft Excel for the geodatabase and contained the following fields: “CAVE_KEY”; “CAVE_LETTER”; “CAVE_NUM”; “CAVE_CODE”; “CAVE_NAME”; “point present”; “Georeferenced”; “Digitized”; “Map available?”; and “Location if available”. The matrix was used to track both georeferencing/digitization and known issues with the data. Information on 392 cave locations was entered into the geodatabase and it, along with the tracking matix, has been delivered to GRCA.

An important issue for the current version of the cave geodatabase is the finalization of primary key fields. Currently, the primary key is a alpha-numerical sequence using a “-“ to separate primary elements within a single column. A more robust method would be to break out cave designator (typically the letter “C”) from the cave number (five digits integer) such that each can then be programmatically concatenated to a third field that contains the entire

key. This would allow for consistent application of the primary key and subsequent joins between the database and geodatabase.