

**GICC Archival and Long Term Access Committee**  
**Final Report Outline – DRAFT v3**  
**May 2, 2008**

**Background –**

- The Archival and Long Term Access Committee was established by the GICC in November 2007 as part of the proceedings for adoption of the GICC Data Sharing Guidelines.
- Recommendations #8 in the Data Sharing Guidelines report included an action on Archive and Long Term Access to geospatial data.
- Specifically, the report recommended that “Data producers evaluate and publish their long term access, retention, and archival strategies for historic data.
- At the GICC meeting, it was noted that the interest in archival and long term access guidelines requires more work, and therefore, the Archival Committee was chartered.
- The Archival committee was seated in February. Ms Anne Payne, Council Member from Wake County GIS, was named Chair. A roster of all members and staff is included on page \_\_\_\_.
- The first meeting of the committee was held on February 29, 2008 at the USGS offices in Raleigh. A list of all meetings held is included on page \_\_\_\_\_

**General Description of the Issues**

- Production points for geospatial data are diffuse which pose many challenges to archiving
- Update cycles for data resources are quite variable, ranging from “daily’ updates for very transactional themes (parcels, addresses) to annual updates (imagery ). The community at-large needs better guidance on desirable frequency of capture schedules for retention and archival purposes.
- Retention and preservation requirements, and schedules, if they exist, are not considered, nor included in up front data life cycle planning, budgeting, nor in work flow development, by local and state agencies.
- Most current documentation practices (metadata) are not adequate for data retention nor archival purposes.

- Long term preservation of geospatial resources will involve some migration and retention of critical documentation.
- The value of historic geospatial data in accessible form has not yet been clearly articulated. Data users are accustomed to working primarily with current data, and, for lack of availability, have not yet discovered meaningful scenarios to utilize historic geospatial records and incorporate analysis into business activities. This scenario is perpetuated by data producers that are overly focused on current data, and as a result, overwrite data frequently.

The Value in Temporal/Historical Data is not yet obvious

- There are emerging examples/cases: shoreline change, land use change, legal proceedings, permit violations, impervious surfaces
- Addressing digitally borne records as the “only” rendition of a record has not yet become incorporated in agency administrative proceedings and planning. (Ex: Imaging analog documents commands higher attention)

See case studies in Appendix \_\_\_\_\_

- The business case has not yet been made for the investment justifications necessary to successfully carry-out a digital geo-preservation plan.
- Users are more likely to discover the value of historic data for business purposes if data can be relied upon and accessible on-line.
- Data that tend to survive are those that are maintained in an on-line service.
- Although data preservation is an emerging area of interest, it is a low priority topic. Opportunities to engage the community and the industry in substantive discussion is desirable. The outreach resulting from the NDIIPP partnership has been beneficial.
- The Existing Risks to Preservation of Digital Geospatial Data
  - Primary Points of Failure
    - Data is not saved, or ...
    - can't be found, or ...
    - media is obsolete, or ...
    - media is corrupt, or ...
    - format is obsolete, or ...
    - file is corrupt, or ...
    - meaning is lost
  - Other Factors
    - Future support of data formats in question, there is no open, supported format for vector data

- Shift to web services-based access, data becoming more ephemeral, data does not get transferred when information can be gleaned or used through portals and viewers
- Inadequate or nonexistent metadata, Impedes discovery and use
- Increasing use of spatial databases for data management

## **On-Going Partnerships on Digital Preservation in NC**

### North Carolina Geospatial Data Archival Project

The North Carolina Geospatial Data Archival Project is Partnership between the NCSU Libraries and NCCGIA, with Library of Congress under the National Digital Information Infrastructure and Preservation Program (NDIIPP). It is one of 8 initial NDIIPP partnerships. A few key elements of the project:

- A focus on state and local geospatial content in North Carolina;
- It is tied to the NC OneMap initiative, which provides for seamless access to data, metadata, and inventories;
- An objective is to engage and leverage NC OneMap collaboration and existing state/federal geospatial data infrastructures in preservation
- Local Government Frequency of capture surveys (2006; 2008)
- Involve NC State Archives agency in the extension work 2008-2009
- Goals:
  - Capture at-risk data for repository
  - Explore technical and organizational challenges for repository
  - Improve temporal data management practices of data producers
  - More efficient means of acquiring and preserving data for archival;
  - Progress towards best practices, including for long term access
  - Outreach and socialization of the problem

### Multi-State Geospatial Content Transfer and Archival Demonstration and Learning Project

This project is a partnership among Archives, Libraries, and GIS state government agencies from Utah, Kentucky and North Carolina. Project is co-lead by NC Archives and CGIA, under the National Digital Information Infrastructure and Preservation Program (NDIIPP). NCSU Libraries is a key partner in the project.

A few key elements of the project:

- Demonstrate and Document Best Practices for:
  - State-to-state transfer of geospatial content using spatial data infrastructure
  - Replication of content
  - Strategies to enable long term access and preservation of geospatial content

- Key Issues
  - Content Selection (which layers needed)
  - Inventory of geospatial content
  - Integrating geospatial content into archival systems (process and frequency)
  - Managing/tracking content once archived
  - Flow of content between states
  - Role of Metadata in the above processes

### **DCR and ERC and Archives Processes**

(This section best if written by Ed, Kelly, Megan, Tom)

### **Guiding Principles and Scope of Work**

*The committee agreed to use the following to guide the development of its recommendations to the GICC:*

- Recommended practices should not place an undue additional workload on state and local GIS professionals. Retention strategies should be easy to accomplish as part of the agencies' normal workflow.
- An organized and structured approach for life cycle creation, management and sharing of geospatial content brings order and efficiencies to the retention and archival process.
- Technical approaches recommended should be designed to minimize the risk of loss of data over time.
- Archiving practices should be consistent with all other GICC-approved standards and recommendations. (Data Sharing Recommendations, Security Guidelines, etc.)
- Recommendations should be consistent with electronic records guidelines, policies and requirements published by the NC Archives – Archives and Records Section.
- Existing retention policies and schedules of local and state agencies should be considered in the development of recommendations
- Existing infrastructure should be employed as much as possible (example - NC One Map Inventory)
- Recommendations should address the following issues:

## DRAFT v3

- What? What content should be preserved? (Keep framework layers in mind)
- When? How often should we create data archives?
- Where? Where are the archived data stored and will they be accessible?
- How? What data formats, compression formats, and media? Should joined attribute data be included?
- Who? Who should be responsible for creation and long-term storage of archived data?

Report's scope and what it *does not* include:

Intended as a beginning guide for initiating data retention and archiving practices; not a comprehensive set of standards. There are recommendations for GIS Content Providers and for Archivists.

Initial guidelines may not equal the rigor of NCDCCR recommendations, but will reflect the spirit and intent of those guidelines

### ***General Best Practices***

#### Archiving Schedule

Should be different for different layers based on frequency of update – Minimum annually for ALL

Schedule for retention may occur on a variable scale...ex; weekly captures for first year or two, retain only an annual version when it is 2 – 10 years old; and then transmit permanent transfer to official long term access site (archives?) or destroy.

#### Inventory

By participating in Inventory Systems (ie NC OneMap Inventory) producers provide data availability information, contact names, minimal metadata, information on rights, technical environment, and in some cases future development/maintenance plans.

#### Storage Medium

Strongly recommend disk drive storage due to preservation advantages – backups, media currency, off-site storage, etc. – IT folks should have a plan in place for this

NCDCCR recommends 3 – 4 copies - preservation master, access record, and backup

#### Formats

Shape files (zipped and non-zipped) At least on unzipped version annually; considered “open source”?

## **DRAFT v3**

### Responsibility

Determine responsible party for long term accessibility, long term preservation

### Metadata Issues

Archives contain metadata; should retention schedules imbedded in metadata?

How and where should details of retention and archiving policies be documented?

### Distribution Availability

OneMap, other?

### Periodic review of policies, data integrity

Check Sums; annual policy review including format media, frequency

NCDCR recommends that you always work from a copy when migrating or making changes to mitigate the risk of data loss

## ***Specific Content Practices***

### Local Primary:

Cadastral w/minimum attributes (specify or suggest?)

Streets Centerlines (suggest attributes here?)

Administrative boundaries – city limits, planning jurisdiction, other?

Orthophotography (mention scanned imagery?)

### Local Secondary:

Political

Zoning

Other?

## **Conclusions and plan for the future**

## Committee Members

Name	Organization
Anne Payne, Chair	Wake County
<a href="#">Kathryn Clifton</a>	City of Salisbury
<a href="#">Mark Crane</a>	USDA
<a href="#">Kelly Eubank</a>	NC Dept of Cultural Resources
<a href="#">John Gallimore</a>	Davie County
<a href="#">Tracey Glover</a>	City of Fayetteville
<a href="#">Amy Keyworth</a>	NC DENR - Water Quality
<a href="#">Bill Lefurgy</a>	Library of Congress
<a href="#">Butch Lazorchak</a>	Library of Congress
<a href="#">Scott Miller</a>	Western Piedmont Council of Government
<a href="#">Tom Morgan</a>	Secretary of State - Land Records
<a href="#">Steve Morris</a>	NCSU Libraries
<a href="#">Zsolt Nagy</a>	CGIA
<a href="#">Doug Newcomb</a>	US Fish and Wildlife Service
<a href="#">Thomas Parrish</a>	NC Dept of Cultural Resources
<a href="#">Joe Sewash</a>	CGIA
<a href="#">Ed Southern</a>	NC Dept of Cultural Resources
<a href="#">Rebecca Troutman</a>	NC Association of County Commissioners

## Committee Meetings

**February 29, 2008**

**May 2, 2008**

## Appendices

Anecdotal/Case Study Support for Preservations  
Related Docs