



Data Preservation Special Interest Group Meeting

Steve Morris
North Carolina State University Libraries

Joe Sewash
NC Center for Geographic Information & Analysis

How would you describe your geospatial archive?



- A folder on your hard drive?
- CDs or DVDs stashed in a desk drawer?
- Hope data was captured during the nightly backup?
- “Old data/ projects” folders on the team file server?
- Integrated with current datasets in your enterprise GIS?
- A formal online archives storage environment ?

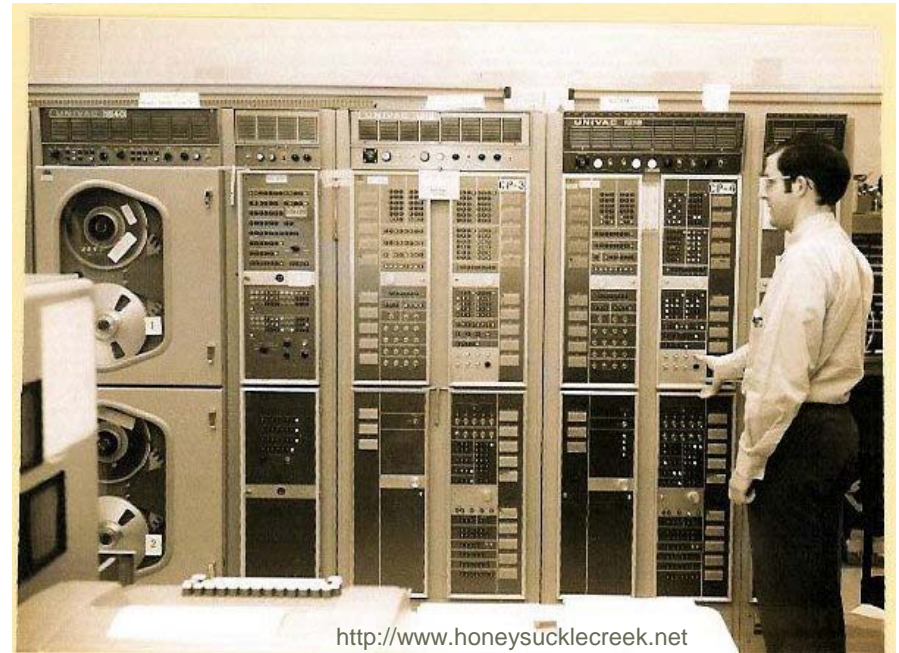
We're doing backups, isn't that archiving?

- ▶ Backups – a means to save and recover current records
- ▶ Personal Archives- “Keeping stuff” on external media or on hard drives or SAN
- ▶ True Archiving – formally preserving important data **permanently** in a trusted digital repository



Digital preservation 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- ▶ media is obsolete, or ...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- ▶ format is obsolete, or ...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



What are the special risks to geospatial information?

- ▶ Unique geospatial data formats
- ▶ Spatial database complexity
- ▶ Fragility and uncertainty surrounding digital cartographic representation
- ▶ Issues related to time-versioned content
- ▶ Metadata unavailability or inconsistency
- ▶ No generally supported content packaging design for complex geospatial data

Library of Congress National Digital Information Infrastructure and Preservation Program (NDIIPP)

Mission: To ensure access over time to a rich body of digital content through the establishment of a national network of partners committed to selecting, collecting and preserving at-risk digital information

- Learn By Doing
- Support Collaboration
- Catalyze Activity
- Break Down Boundaries

NDIIPP Geospatial Projects

National Geospatial Digital Archive
(2004-2009)

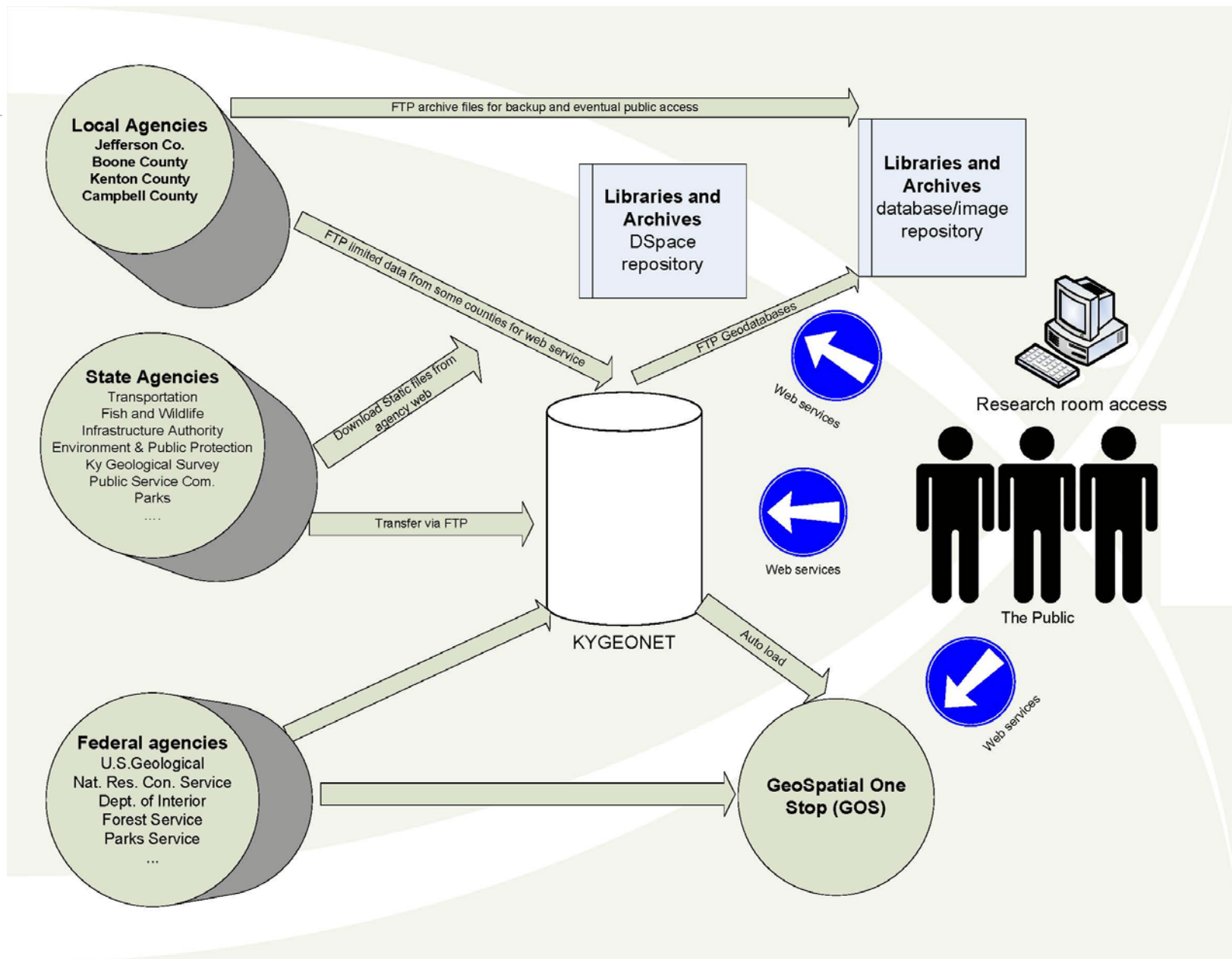
Geospatial Multistate Archive and
Preservation Partnership (2007-2011)

NC Geospatial Data Archiving Project
(2004-2009)

CIESIN Geospatial Preservation Portal
(2010-)

What is GeoMAPP?

- ▶ Preservation of “at risk” superseded geospatial content
- ▶ Building the relationship between State GIS and Archives staff
- ▶ Interstate partnerships (NC, KY, UT and new partners)
- ▶ Exploration of business case drivers
- ▶ Data replication among several states
- ▶ Implementation of a geographically dispersed content-exchange network
- ▶ Local, state and national outreach



Big picture GeoMAPP takeaways

- ▶ Collaboration is a key
 - ▶ GIS & Archives staff
 - ▶ Get to know your data producers
- ▶ Know what you have (data inventory)
- ▶ Make it official (data appraisal/ records scheduling)
- ▶ Leverage existing workflows and investigate new sustainable processes to make the data last
 - ▶ Don't re-invent the wheel
 - ▶ Keep data discoverable/ accessible/ usable for future use
 - ▶ Justify the investment (business case)



Questions for discussion

- ▶ What is the level of awareness in the vendor community?
- ▶ Spatial databases (Geodatabases): How are you handling temporal data management and archiving?
- ▶ How you handling archiving of GIS projects?
- ▶ How are you handling capture and archiving of cartographic representations?
- ▶ What approach do you use for transfer of data between organizational entities?
- ▶ How are you handling storage for large raster collections?
- ▶ If you are capturing snapshots of vector data, with what frequency do you capture versions?

Question: What is the level of awareness in the vendor community?

- ▶ Is archival support on the radar for current / future feature development in products?
- ▶ Are customers requesting more support from vendors to support archival features in future releases?
- ▶ What are the most effective mechanisms for clients and vendors to collaborate on developing archival support features?

Question: How are you handling spatial databases?

- ▶ Managing temporal data forward in active database?
- ▶ Transferring database snapshots to an archive?
- ▶ Transactional updates to an archive?
- ▶ Extracting feature datasets as shapefiles?

Contacts

- ▶ Steve Morris (NCSU)
steven_morris@ncsu.edu

- ▶ Joe Sewash (NCCGIA)
joe.sewash@its.nc.gov

<http://www.geomapp.net>

