The Earth Institute at Columbia University

Mobilizing the Sciences and Public Policy to Build a Prosperous and Sustainable Future

World Data Center for Human Interactions in the Environment

Practices for Managing and Preserving Geospatial Data

Robert R. Downs, Robert S. Chen,
Mark Becker, John Mickelson, John Scialdone
Center for International Earth Science Information Network
Columbia University
www.ciesin.columbia.edu

19th Annual Northeast ARC Users Group Conference
Lake Placid, NY
September 21, 2004

The project
"Managing and Preserving Geospatial Electronic Records"
is investigating the requirements for state, county, and municipal agencies to manage and preserve geospatial data and related electronic records, especially those generated by GIS software.

- Funded by the National Historical Publications and Records Commission (NHPRC) of the National Archives and Records Administration (NARA).

The project seeks to identify and recommend practical and appropriate policies, techniques, standards, and practices

- to manage geospatial electronic records (GERs)
- to support their long-term retention and dissemination and
- to facilitate their usability and utility as important information resources of significant historical interest

Major Tasks of Project

- Establish an Interdisciplinary Advisory Board from Region
- Assess Needs for Managing Geospatial Data and Records
- Review Current Research and Relevant Standards
- Design Data Model to Manage and Preserve Geospatial Data
- Develop Resource Guide and Instructional Materials
- Conduct a Workshop for Agency Practitioners and Managers
- Disseminate Guidance Materials on Web Portal
- Publication of a Research Paper Summarizing Study

Project Advisory Board and Workshop Program Committee

- **Cheryl Benjamin**, Standards/Data Coordination NY State Cyber Security & Critical Infrastructure Coordination
- **Kevin Glick**, Electronic Records Archivist, Yale Univ. Sterling Memorial Library
- **Bill Guthe**, Educational Tech Center GIS Analyst, Princeton University
- **Geoffrey Huth**, Mgr, Records Services Development NY State Education Dept, Archives
- **Sally Johnson**, Policy Analyst for Statewide Planning Program, State of Rhode Island
- **Patrick McGlamery**, Library Liaison to Geography, University of Connecticut
- **Jennifer O'Neill**, E-Government & GIS Specialist, NY State Education Dept, Archives
- **Theresa Pardo**, Deputy Director, Center for Technology in Government, University of Albany
- **Robert Sandev**, GIS Officer, Division for Ocean Affairs and the Law of the Sea, Office of Legal Affairs, United Nations
- **Doug Schuetz**, GIS Director, Rockland County Planning
- **Bridget Sisk**, Chief, Archives & Records Mgt Section, United Nations Secretariat
- **Terri Spies**, GIS Specialist, NYC Dept of Environmental Protection
- **Larry Thornton**, Bureau Chief, Geographic Information and Analysis, NJ Dept. of Environmental Protection
- **Sam Wear**, GIS Director, Westchester County

The needs assessment study takes a qualitative approach: Interviewing professionals to assess current and future needs of government agencies for managing and preserving geospatial data and related electronic records.

Needs Assessment Study: Obtaining a Cross-Section of Perspectives and Experience from a Diverse Group of Professionals

- Representation: Primarily the Northeastern Region of the U. S.
- Fields: Planners, Geographers, Systems, Librarians, Archivists
- Responsibilities: Directors, Managers, Analysts, Developers
- Organizational Types: Government, Education, Non-Profit
- Government Levels: Federal, State, County, Local

Needs Assessment Study: Categories of Initial Observations

- Data Distribution Challenges
  - Security-Related Constraints for Internal and External Distribution
  - Protecting Sensitive and Confidential Data
  - Liability, Intellectual Property, and Digital Rights Management
- External Cooperation
  - Partnerships and Inter-Organizational Cooperation
  - Community Resources
- Internal Coordination
  - Data Development and Organizational Coordination
  - Quality Assurance and Documentation
  - Archival and Preservation Needs
- System Capabilities
  - System Security and Risk Management
  - Data Volume

Examples of Observed Needs for Data Distribution Challenges

- Control distribution of data distributed internally to prevent external access
- Users prohibited from re-disseminating data
- Sanitize data or restrict access to entire dataset for any confidentiality concerns
- Identify and record Intellectual Property rights, citation constraints, and restrictions for access and use when acquiring data

Examples of Observed Needs for External Cooperation

- Maintain stock written agreements for data sharing
- Form web services partnerships to provide access to and integrate with data served by others
- Develop library of case examples, best practices, and recommended technologies for geospatial data storage and delivery
- Establish peer-review council to review data to be released for public use or included in clearinghouses
- Organizations need to share lessons learned on implementing digital repositories

Examples of Observed Needs for Internal Coordination

- Budget to maintain data of continuing value initially acquired for previous programs
- Document techniques developed, assumptions, conclusions, and methods for integrating datasets
- Implement a centralized digital repository system to manage data and metadata
- Integrate geospatial data with records management program
- Layer by layer committee review for dissemination and archiving
- Establish minimum documentation requirements and templates
- Allocate sufficient staff and cross-train to manage data
- Archive data on CD and DVD media and store one copy off site

Examples of Observed Needs for System Capabilities

- Maintain software and hardware to access old backup copies of system
- Establish database management system to allow personnel to track data
- Separate external and internal database management systems for independent performance
- Establish capabilities to control access to layers or parts of data
- Prepare for increasing data volumes with large storage capacity servers and high bandwidth networks
- Separate servers storing internally and externally accessible data
- Require login passwords to access any machines on internal network
- Assign “write” privileges for data repository only to data librarians
- Deposit tape backups to onsite and offsite storage

Key Standards Being Reviewed

- Content Standard for Digital Geospatial Metadata (CSDGM), Federal Geographic Data Committee (FGDC). FGDC-STD-001-1998 http://www.fgdc.gov/metadata/contstan.html
- Geographic Information Metadata (ISO 19115)
- Reference Model for an Open Archival Information System (OAIS) Adopted as ISO 14721:2003 http://www.ccsds.org/documents/650x0b1.pdf
- Producer-Archive Interface Methodology Abstract CCSDS 651.0-R-1 http://ssdoo.gsfc.nasa.gov/nost/isoas/CCSDS-651.0-R-1-draft.pdf
- NARA Records Management. Expanding Acceptable Transfer Requirements: Transfer Instructions for Permanent Electronic Records, Digital Geospatial Data Records.
- NARA Code of Federal Regulations Part 1234 Electronic Records Management (5/16/01).
- DOD 5015.2-STD (June 19, 2002) C2.T3 Record Metadata & C4.T1 Classified Record Components. http://jitc.fhu.disa.mil/recmgt/standards.htm
- Information and Documentation – The Dublin Core Metadata Element Set (ISO15836) http://www.niso.org/international/SC4/n515.pdf

Extending the Boundaries: A Workshop on Managing and Preserving Geospatial Electronic Records

Palisades, NY
May 13, 2004

"Extending the Boundaries: A Workshop on Managing and Preserving Geospatial Electronic Records", was convened to engage practitioners and managers from both the geospatial and data archiving communities in collaborative discussions that address the challenges faced by state and local government personnel in managing and preserving geospatial data and related electronic records.

Workshop Participants

- 25 Professionals from NY, NJ, CT, and RI
- International, Federal, State, County and Local Agencies and Educational Institutions Represented
- Diversified Expertise
  - Geospatial Project Management and Coordination
  - Geographic Product Analysis and Development
  - Planning and Policy
  - Archives
  - Geographical Data Coordination and Libraries

Themes of Workshop Breakout Sessions

Coordination and Infrastructure for Sharing
Security, Confidentiality, and Freedom of Information
Preservation for Future Access and Use
Quality Assurance, Documentation, and Reusability

Workshop Results: Themes of Recommendations

Theme	Description
Policy and Implementation	Policies provide direction and guidance for agencies to initiate and implement change.
Climate for Change	The organizational culture needs to value change and promote opportunities for improvement.
Outreach Coordination	Share resources to improve practices and facilitate learning among staff within several agencies.
Communication Among Archives/GIS	Share knowledge and promote learning between work groups, disciplines, and organizations.
Justification Examples	Demonstrate the benefits and costs of managing geospatial data to justify the allocation of needed resources.
Technical Standards and Interoperability	Proven practices, tools, and techniques for managing and preserving geospatial data.

Workshop Recommendations: Policy and Implementation

- Develop criteria for identifying archival records in GIS
- Create a good standard and emergency data/record distribution policy
- Identify secure and confidential data/record based on Freedom of Information
- Create data sharing agreements with all parties (sooner, not later)
- Protection of sensitive/confidential records by destruction is not an option; Cannot destroy records as a way to protect them
- Create secure and confidential data distribution mechanism with secure access controls
- Continuously monitor changes to security/privacy of records throughout the lifecycle of the records (Communication between records creators and preservationists)

Workshop Recommendations: Climate for Change

- New state law to fund and define mission for archives and records management of GIS – for those states without.
- Think outside the GIS/Archiving box (Legal, HR, IT, Project Management, PR, Legislative, Users)
- Integrate activities into existing initiatives so they do not seem to be new work
- Create an enforcement mechanism for each distribution policy
- Involve archivists/records professional in the creation of secure/confidential data distribution mechanisms that produce preservable records (including access logs)
- Create an organization with joint mission: grab turf

Workshop Recommendations: Outreach Coordination

- Continue the work of this MAPGER project
- Web portals for outreach, education, and curricula
- Website of case studies on how to integrate GIS and archives
- Grant program to develop a GIS/archival consortium

Workshop Recommendations: Justification Examples

- Share case studies with cost savings and public benefit
- Centralizing and pooling interagency cooperation vertically
- Fundable project plans (include ROI for groups)
- Put money and resources into distribution
- Manager allocate resources to preserve GIS records

Workshop Recommendations: Communication Among Archives and GIS

- Joint workshops among professional groups
- Provide training and certification program for metadata and archives
- Cross-training of archivists and the GIS field is necessary: archivists to preserve GIS, GIS to produce preservable records, create archivist-type position within GIS organization
- GIS and archivists participate in committee to select GIS records for preservation
- Training on standards for archiving GIS records and techniques for using technology

Workshop Recommendations: Technical Standards and Interoperability

- Promulgate open GIS standards and internet map coordination
- Open source development of standards for archival and GIS records community
- Develop an open non-proprietary GIS format

Visit the project website for future updates:

Managing and Preserving Geospatial Electronic Records

www.ciesin.columbia.edu/ger

Thank you!

Rdowns@ciesin.columbia.edu