



Library of Congress Digital Preservation Newsletter

NDIIPP Sponsored GIS Tutorial



Downtown Pittsboro, North Carolina, U.S.A., from the 1997 Chatham County North Carolina digital orthophotography capture.

Numerous points social and physical information are tracked over time with geospatial data, like the erosion of a coast line or the shape of commercial development in a particular area. This specialized data is used by national and local governments and is an important part of the public record. The Library of Congress is supporting several efforts to preserve geospatial data and on April 21 sponsored a day-long tutorial called “Preservation Issues Related to Digital Geospatial Data.”

The tutorial, led by [Steve Morris](#), head of Digital Library Initiatives at North Carolina State University Libraries and co-principal investigator of the NDIIPP-supported North Carolina Geospatial Data Archiving [Project](#), provided an overview of the process of creating geospatial data, how it is used by different groups, and the different strategies for making the data accessible over time. “GIS can be used in a wide range of local government appli-

To subscribe to this newsletter, go to http://service.govdelivery.com/service/subscribe.html?code=USLOC_13, type in your e-mail address, scroll down and click on “Digital Preservation.”

cations such as economic development, tax assessment, management of public utilities and emergency response,” said Morris. “Data produced at the local level can be supplemented with data from state and federal agency sources.”

The tutorial was attended by staff across many divisions at the Library of Congress and several other government agencies, including: the Congressional Research Service, the U.S. Geological Survey, the National Archives and the Government Printing Office.

For more information about the talk, visit the NDIIPP [Web site](#) and stay tuned for information about more NDIIPP sponsored tutorials. ■

BagIt: A New Transfer Format

The Library of Congress and the California Digital Library have jointly developed a format for transferring digital content. “BagIt” is based on the concept of “bag it and tag it,” where digital content is packaged (the bag) along with a small amount of machine-readable text (the tag) to help automate the content’s receipt, storage and retrieval. There is no software to install.

BagIt is an attempt to simplify large-scale data transfers between cultural institutions. It builds on the success of the Library’s data-transfer projects, including the [Archive Ingest and Handling Test](#) and the [San Diego Supercomputer Center data-storage project](#). BagIt streamlines the process and reduces the number of “moving parts.”

Though data transfer can be done by network or disk, the Library expects to receive more and more digital collections over the network. Conforming to the BagIt format will guarantee users a smooth, orderly transfer, with little or no human intervention. John Kunze, Preservation Technologies Architect at the California Digital Library and one of the principle authors of BagIt, said, “It really makes it easy for another institution to hold onto some of our content, and that’s reassuring.”

A “bag” consists of a base directory, or folder, containing the “tag” and a subdirectory that holds the content files. The tag is a simple text-file manifest,



like a packing slip, that consists of two elements:

1. An inventory of the content files in the bag
2. A checksum for each file.

A checksum is a way to validate that everything in the bag arrived OK. Once the content is validated, the receiver notifies the sender, usually by e-mailing a confirmation receipt.

Each line contains a string of characters – the checksum – followed by the name of the file and its directory path. See the example below:

```
408ad21d50cef31da4df6d9ed81b01a7 data/docs/exmpl.doc
```

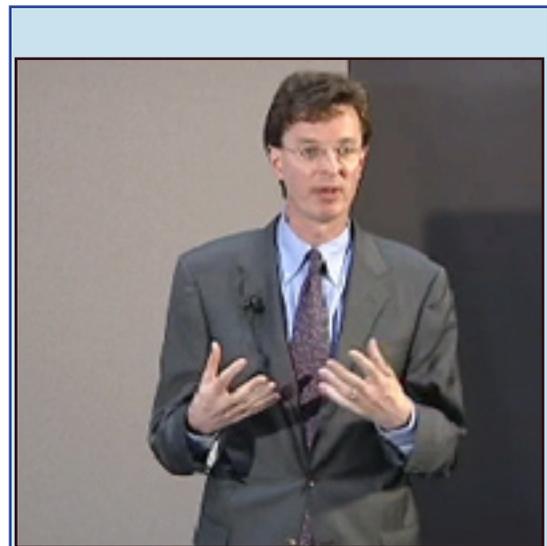
In another optional file, users can add metadata about the content, such as a description of the package contents and detailed contact information for the sender.

A slightly more sophisticated bag lists URLs instead of the actual content files. A script then consults the tag and fetches the desired files over the Internet, ten or more at a time. This type of simultaneous multiple transfer greatly reduces the overall data-transfer time

The full BagIt specification is available at: <http://www.cdlib.org/inside/diglib/bagit/bagitspec.html> ■

Digital Preservation Outreach

The Library of Congress has developed a short flash presentation to help visitors to digitalpreservation.gov learn more about why digital preservation is important and how the Library and its partners are developing a national digital preservation plan. This short presentation is available [online](#), and it has been exhibited at the National Bookfestival on the National Mall in Washington, DC. ■



Robert Horton, director of Library, Publication, and Collection at the Minnesota Historical Society

Meeting Gets State Legislators Thinking About Digital Preservation

Last December, the Library's NDIIPP Program made four awards to projects working with 23 states to collect and preserve state and local digital content at risk of loss. The Minnesota Historical Society leads one of the "States" projects and were joined by members from the National Conference of State Legislatures from 12 states to [kick off their work](#).

The project is working with legislatures in California, Kansas, Tennessee, Mississippi, Illinois and Vermont, to explore improved and enhanced access to legislative digital records. This will involve implementing a trustworthy information management system and testing the capacity of different states to adopt the system for their own use. Content will include bills, committee reports, floor proceedings and other legislative materials.

Robert Horton, director of Library, Publication, and Collection at the Minnesota Historical Society noted "that use value is going to drive the functions" of the information system, and a number of participants echoed the idea that improved access to the legislative materials would lead to increased use, which would help support the case for preservation. ■