

Monday, April 3, 2006 3:30 – 5:00

Background of Presentation:

The presentation outlined below discussed CD-ROM Data Migration, as presented by Lisa Russell, Content Manager, Planning and Development Content Management. This served as one portion of the Council Session entitled, “GPO Digital Content Forum.” Facilitated by Richard Davis, Director, Library Services and Content Management, this session offered four separate presentations on:

- **Web Harvesting**
- **CD-ROM Migration**
- **Digitized Content and**
- **Digitized Content Specifications**

CD-ROM Data Migration

Key Discussion Points - GPO is developing a strategy to maintain access to the content of depository CD-ROMs over time to ensure permanent public access to discs that utilize older technology.

- Depository libraries currently attempt to maintain older computers in order to access government information content on older formats that may not run on newer technologies.
- GPO is addressing this problem by evaluating options for preserving access to the content of older CD-ROMs. Other media are also issues, but GPO is starting with CD-ROMs to start to devise a strategy and will address the other media at a later date.
- Addressing at-risk content that may be lost is the first priority.
- Data preservation may include refreshing data, data migration, and/or emulation strategies.

Summary:

- As technology changes, it becomes more difficult to access content on older media.
 - The hardware and software used for these products become obsolete, and it becomes difficult to maintain older computers needed to access these products.

- Proprietary software requires maintenance of software licenses or may not be supported by vendors as they develop newer versions of the software, making it difficult to access content on older products. Some software vendors have gone out of business, leaving GPO and depository libraries without technical support for their products.
- GPO is taking steps to address the problem by identifying issues, identifying potential solutions, evaluating potential solutions, developing a strategy, and implementing a solution.
- Potential solutions include refreshing data, data migration, and emulation technology.
 - Refreshing data is moving data to new storage media periodically.
 - Data migration is the periodic transformation of files into a new digital format, ensuring compatibility between file formats and applications.
 - Emulation allows obsolete systems to be run on future systems, maintaining the use and application of the original documents with the original software.
 - A final strategy may include some or all of these solutions to ensure access to content.
- GPO staff conducted a review of a sample of CD-ROMs included in an inventory spreadsheet provided by the Regional depository at the University of Kentucky.
 - Three agencies were used for the sample: Department of Education, Department of Justice, and United States Geological Survey (USGS).
 - Discs were prioritized based on whether the content was found on the Internet in order to ensure that content that was not available elsewhere was addressed first.
 - Lessons learned from the sample will be used to develop a strategy for handling discs from other agencies.
- GPO is working with other stakeholders to address data migration issues.
 - GPO is an active participant in the Government Information Preservation Working Group (GIPWoG). (See <http://www.itl.nist.gov/div895/gipwoq/index.html> for more information on GIPWoG.)
 - GPO is working with the National Institute on Standards and Technology (NIST) on a pilot project to test emulation technology as a potential solution to make older discs accessible.
- There was a brief demonstration of the emulation pilot that is GPO is conducting with NIST.