

Depository CD-ROMs: Maintaining Access to Content as Technology Changes

Lisa Russell

Content Manager

Planning & Development Content Management

US Government Printing Office

April 3, 2006



U.S. GOVERNMENT
PRINTING OFFICE
KEEPING AMERICA INFORMED

Depository CD-ROMs

- ◆ During the fiscal years 1995 through 2002, 4901 tangible electronic products were distributed to FDLP libraries
- ◆ Media has changed over time, as seen in Library Program Service Annual Reports
 - FY1996 statistics on “tangible electronic” titles
 - FY1998 statistics on “CD-ROM” titles
 - FY2001 statistics on “CD-ROM, DVD” titles



Problem

- ◆ How to ensure that the government content on tangible electronic media that have been distributed to FDLP libraries will continue to be accessible?
 - Hardware and software become obsolete
 - Difficult to maintain older machines
 - Some use proprietary software



Steps to Solving the Problem

- ◆ Identify issues
- ◆ Identify potential solutions
- ◆ Evaluate potential solutions
- ◆ Develop a strategy
- ◆ Implement solution



Identifying Issues

- ◆ Deal with CD-ROMs first, other media later
- ◆ Start with a sample to acquire data on issues
 - Inventory spreadsheet from University of Kentucky
 - Selected 3 agencies to use as a sample
 - Department of Education
 - Department of Justice
 - US Geological Survey (USGS)



Identification of Sample

- ◆ University of Kentucky inventory
 - Conducted by Tom Byrnes
 - Began June 2003
 - Completed May 2004
- ◆ 3398 discs listed in spreadsheet
 - 69 discs from Department of Education
 - 31 discs from Department of Justice
 - 748 discs from USGS
 - 848 discs in sample



Starting the Analysis

- ◆ Establish priority levels for dealing with discs
- ◆ Search for Internet versions
- ◆ Content first, functionality later
- ◆ Use lessons learned from sample to devise strategy for dealing with all discs



Prioritization

- ◆ Three priority categories emerged in sample
 - Most Critical
 - Titles not located on the Internet considered to be at greatest risk of losing content
 - Intermediate
 - Some titles found on the Internet, but require more investigation to determine if they are complete
 - Considered more critical than titles that are complete on the internet, but less critical than titles that are not on the Internet at all
 - Less Critical
 - Some titles less functional on the Internet, but less risk of content loss than titles not on the Internet



Components of Data Preservation

- ◆ Refreshing data
 - Moving files to new storage media periodically
- ◆ Data migration
 - Periodic transformation of files to 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 original software



Partners

- ◆ Government Information Preservation Working Group (GIPWoG)
 - Formed in 2003 by a group of agencies interested in data preservation
 - Mission: To provide government agencies with information to make informed decisions for using digital storage technologies in digital data storage or preservation strategies
 - <http://www.itl.nist.gov/div895/gipwog/index.html>



Partners

- ◆ National Institute for Standards and Technology (NIST)
- ◆ July 2005
 - Meeting at GPO to discuss preservation issues
 - Copies of 23 titles given to NIST for testing
- ◆ October 2005
 - Meeting at NIST
 - Wo Chang presented early results of emulation test on discs supplied by GPO



Partners

- ◆ National Institute for Standards and Technology (NIST)
- ◆ January 2006
 - Results of emulation testing presented at GIPWoG meeting held at GPO
 - GPO and NIST agreed to further collaboration on emulation testing
 - GPO provided copies of 5 additional titles that utilize proprietary software



Next Steps

- ◆ Further review of the CD-ROM sample
- ◆ Release list of reviewed CD-ROMs
- ◆ Continue discussions with GIPWoG and other interested parties
- ◆ Continue testing of emulation solution
- ◆ Evaluate potential solutions
- ◆ Devise and implement strategy for permanent public access



Demonstration of Emulation Solution

Lisa Russell

Content Manager

Planning & Development Content Management

US Government Printing Office

April 3, 2006



U.S. GOVERNMENT
PRINTING OFFICE
KEEPING AMERICA INFORMED