



Web Documents Digital Archive Pilot Project at GPO

George Barnum
Depository Library Conference
16 October 2001



Today's session

- Briefly explain the goal for the project
- Explain the background and progress of the project
- Discuss GPO's particular emphases

Background of the project

- Long history of GPO/OCLC cooperation
- PURLs
- FDLP/ERIC Pilot
- Current project in planning since late 1999
- “Build us an archive and the tools to use it”

What will the project accomplish?

- An offsite, vendor-maintained archive
- A “toolkit” for processes
- Defined set of preservation metadata
- Integration of workflow and tools

What kind of “toolkit”?

- Currently, we use a wide variety of tools to acquire, classify, track, distribute, archive, and catalog:
 - ACSIS
 - DDIS
 - PAMALA
 - OCLC Prism
 - Teleport Pro
 - FTP
 - Manual systems

What’s the common denominator?

- ACSIS
- DDIS
- PAMALA
- OCLC Prism
- Teleport Pro
- FTP
- Manual systems
- They help us create, gather, and store data about the publications

What do we do with the data?

- Archive for permanence
- Catalog for bibliographic access
- Track for internal production standards
- Manage the FDLP Electronic Collection

The Goal of the Project

- is to create a set of integrated tools and processes that will help us cope with the ever-increasing number of electronic publications and take the best possible advantage of technology

What will “it” do?

- Initially
 - Gather data out of the acquisition and classification process
 - Track electronic pubs beginning with identification
 - Route pubs to the GPO archive
 - Create preservation metadata
- Eventually
 - Assist discovery (electronic gray bins)
 - Route pubs to multiple archives, including GPO, partners, and OCLC “Digital Vault”
 - Assist in managing pubs in the archive

What will “it” do?

- Provide a platform for various processes:
 - Acquisition
 - Classification
 - Cataloging
 - Archive management

When?

- First phase rolled out in September, 2001
- Phase 2 begins February, 2002

Will it replace systems we use now?

- Not initially. It will work in tandem with current systems and processes. In some cases it will exchange data directly, in others not.

The Plans

- Phase 1:
 - Create metadata
 - Define workflows
 - Test CORC functionality
- Phase 2:
 - Harvest
 - OCLC operated archive server

Questions?