U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

Preservation of Federal Digital Publications

GPO, Federal Depository Libraries, and Other Federal Agencies

Robin Haun-Mohamed & Gil Baldwin

GPO's Mission

- "Perpetual, free, and ready public access."
- Preservation elevated to an agency responsibility
- Future enhancements of GPO's capability in a distributed environment

A Library View of Preservation

- GPO administers the FDLP to promote public access to U.S. Government information
- Permanent access is the goal, preservation is the means

Collaborative Preservation Work

- FDLP Electronic Collection Plan
- NARA Archive Affiliate Agreement
- GPO Access
- Library Partners
- Agency Partners
- OCLC Digital Archive
- Offsite Data Backups

Current Efforts

- Content delivery via *GPO* Access
- Integrated Library System for bibliographic metadata
- Digitizing Federal publications
- Harvesting Federal publications from the Web
- LOCKSS Pilot Project

Collaboration: Early Steps

- Meetings of experts on digitization specifications and metadata for preservation
- Development of white paper, Priorities for Digitization of Legacy Collection
- Published specifications for digitization of Federal publications for preservation

Strategic Vision

- Strategic Vision shared in December 2004
- "To deliver Federal information products and services from a flexible digital platform"

From Vision to Reality

FDsys

GPO's Future Digital System

FDsys: GPO's Future

- Future Digital System (FDsys)
- Will meet the preservation and access goals for all of GPO
- Replaces numerous legacy systems

FDsys from 50,000 Feet

- Goal: Trusted Repository
- OAIS reference model
  - package-based system
  - flexible and interoperable

Critical Function of the FDsys: Preservation

- Protection of the AIP
  - Redundant instances of content
  - Multiple layers and firewalls
- Preservation Processes
  - Refreshment
  - Migration
  - Emulation

Collaboration on FDsys

- External customer input
- Development of ConOps, requirements, and acquisition strategy shared with broad community
- Interagency Council on Digital Content Submission
- Library community
- Collaboration with LC and NARA

FDsys Development

- August 2006: Harris Corp. – GPO's Master Integrator
- March 2007: First release (internal)
- August 2007: Initial public release
- 2008: Full functionality

FDsys Scope

- Content in scope for the FDLP will be preserved
- Content from multiple sources
  - Deposited content
  - Converted content
  - Harvested Content

Harvesting Pilot Project

- Identify and harvest born digital Federal publications
- Contract with Two Companies
- Test Methodology: Crawls of the EPA Web site
- Findings

Digitization Pilot

- First stand-up of FDsys
- Test creation of SIP
- Validation of digitization specifications

Next Steps

- Integration of the ILS
- Migration of legacy content
- Content authentication

Preservation of Federal Digital Publications

Questions?