

# **PRIORITIES FOR DIGITIZATION OF HISTORICAL COLLECTION**

U.S. Government Printing Office

Washington, DC  
July 20, 2009

## 1. Introduction

GPO, working with libraries and others in the information community, has embarked upon a collaborative digitization process by which all Federal publications back to the earliest days of the Federal Government will be made available in a digital format. A digitization project of this magnitude will only be accomplished by prioritizing those publications that should be among the first to be digitized.

The objective is to ensure that the digital materials are available, in the public domain, for no-fee permanent public access through the FDLP. The conversion of tangible materials will begin with print publications, but will eventually include microfiche and other tangible formats. Information will be digitized based upon priorities as identified below or local needs. The digital files and the associated metadata will be preserved in the GPO electronic archive (in addition to any other places that the materials might be held) and there will be no-fee public access to the content through files ingested into FDsys or available on *GPO Access*.

## 2. Background

Digitization of the Federal publications collection will be a very large undertaking with significant staff and financial resource implications. Prioritization of materials to be converted will be essential to effectively manage resources, while serving a variety of other internal and community/user needs. For these reasons, GPO undertook a multi-step approach to determine a priority plan for digitization that included collaboration with other interested parties in the information community. This included a survey in May of 2004 of materials to digitize and the incorporation of input from organizations such as the Association of Research Libraries (ARL) and the Government Document Roundtable's (GODORT) Rare and Endangered Government Publications Committee.

Survey participants were asked to review a composite list as presented and then add other titles that they felt should be included. The results of the survey were then compiled, consolidated and released for the next part of the survey.

Following the analysis and consolidation of the recommendations, GPO asked the community to rank suggested titles and series for digitization. This ranking period for the survey was open from June 23, 2004 through August 6, 2004. The survey process allowed for ranking by depository and non-depository libraries. Each participant was allowed to select and rank up to ten titles for digitization, based on personal experience and perceived local or national need. Real-time ranking results were made available for both depository and non-depository libraries. A total of 310 depository libraries and 134 non-depository libraries participated in the survey.

Of particular interest was the identification of the Congressional Record, U.S. Congressional Serial Set, Statistical Abstract of the United States, and U.S. Statutes at Large in the first four positions for both depository libraries and non-depository libraries. In addition to the survey, members of GODORT's Rare and Endangered Government Publications Committee provided recommendations to GPO on the most important titles to be digitized. Three reviewers recommended one title, the Annual Report of the War Department, as a priority.

In the 2005 white paper *Priorities for Digitization of Legacy Collection*, Library Services and Content Management (LSCM) also worked with other areas of GPO to ensure sufficient material was being identified and obtained for a 2006 scanning project performed in-house by GPO staff as the first public stand up of FDsys. Resources reviewed for the initial white paper include *Report of the Meeting of Experts on Digital Preservation: Digital Preservation Masters*, *Report of the Meeting of Experts on Digital Preservation: Metadata Specifications*, the results of the Digitization Ranking Survey, specific recommendations based on GODORT's Rare and Endangered Publications Committee, a review of use statistics for the databases on *GPO Access*, and feedback from the Fall 2004 breakout sessions for different types of libraries on digitization priorities.

The priority list as shown below has been developed based in part on the review of recommendations and resources conveyed to GPO by librarians in the Federal Depository Library Program and greater information community. In early 2009, the original priority list was reviewed to determine if changes were needed. After reviewing the lists and the digitization projects GPO has participated in, the list essentially remains unchanged. The most useful and most highly used resources are included. The goal of moving systematically back in time from the resources on GPO Access and FDsys allows for a collaborative and orderly process in which to identify and work with others to digitize these important Federal resources.

### 3. Priorities

Code of Federal Regulations	1990-1996
Federal Register (including Unified Agenda)	1990-1994
U.S. Code (index and Supplements)	1988
U.S. Code (index and Supplements)	1994
U.S. Code (index and Supplements)	2000
Congressional Record (Daily Edition)	1989-2003
Congressional Record Index	1989-2003
Bound Congressional Record	1989-2003
U.S. Reports	1990-forward
Public Laws	101 <sup>st</sup> -103rd Cong
Private Laws	101 <sup>st</sup> -103rd Cong
Congressional Bills	101 <sup>st</sup> -102 <sup>nd</sup> Cong

#### 4. Issues to be Addressed:

After the completion of retroactive digitization back to 1990, GPO will reassess digitization priorities. GPO will proceed with collaborative digitization of publications currently located on GPO Access and FDsys in 10-year increments until all preceding publications have been digitized back to publication inception. It may be possible for priorities to conflict with reality of available documents and collaborative agreements sufficient keep workflow moving. In addition, disclaimer statements should be developed to reflect the situation when documents are not available and therefore sets may be incomplete. The plan should be reassessed as needed in response to collaborative community digitization effort.