

**UNITED STATES GOVERNMENT PRINTING OFFICE  
(GPO)**

**CONCEPT OF OPERATIONS DOCUMENT  
CONOPS V1.0**

**FOR**

**COMPOSITION SYSTEM  
REPLACEMENT**

Final  
November 29, 2006

FINAL

## Document Change Control Sheet

**Document Title:** Composition System Replacement (v1.0)

Date	Filename / version #	Author	Revision Description
7/12/2006	CSR Conops v1.0.doc	PMO	Draft Conops, revised from Conops dated May 2006.
7/18/2006	Same	Fleetwood	Formatting, content changes, and Green's input.
7/25/2006	Same	Fleetwood	Added Reeves' content to section 5.5. Additional content changes.
7/25/06	Same	Schwartz, Fleetwood, Knoll	Accepted changes, assigned sections.
8/8/06	Same	Fleetwood	Replaced the original Preface with Green's Preface. Updated the following sections: 3.1.1; 4.4; 8; 8.1; 8.2; 8.3; 3.5.2; 5.5.2; 5.5.4; 6;
8/17/06	Same	Schwartz	
9/8/06	Same	CSR Team	Review of Section 4 Justification for and Nature of Changes.
11/14/06	Same	CSR Team	Updated to reflect Lyle Green's changes sent 11/13.
11/17/06	Same	PMO	Last edits prior to external review. Changed version to v1.0.
11/28/06	Same	Knoll/Fleetwood	Incorporated IT&S and G. Barnum's recommended changes.

**FINAL**

## **Preface**

The United States Government Printing Office's (GPO) current composition system is based on a 30-year-old batch composition engine, developed and maintained by GPO, called Microcomp. Microcomp is currently used to compose the majority of Congressional documents and select Federal agency publications that are printed and published electronically by GPO.

Microcomp was originally designed to compose data files with typesetting-specific codes called locators. Locator codes are tied to external format files that specify style attributes to be applied to segments of content. Over the years, approximately 700 related applications and utilities (e.g., translation tools, delivery tools) have been developed to enhance and sustain the Microcomp composition process in response to the evolving needs of GPO's Congressional customers and changing in-house print and electronic access requirements.

In response to Congress' adoption of XML as a data standard, GPO attempted to retrofit Microcomp to compose Standardized General Markup Language (SGML) / Extensible Markup Language (XML). Subsequently, Congress sought to improve workflow processes for content creators as well as search and retrieval of data by end users. As a near-term solution, the retrofitting of Microcomp was suspended and GPO decided instead to translate XML files into locator codes for printed publications.

As a long-term strategic decision, GPO has decided to replace the system altogether, a project referred to herein as Composition System Replacement (CSR). The envisioned system will work within GPO's enterprise architecture (EA) as well as operating as a stand-alone application deployed remotely at customer sites. The system will be rules-based to facilitate the migration of content into future data tagging schemes/technologies while utilizing industry best practices.

CSR must match the typographic style and page layout of current printed publications as well as support enhanced search, retrieval, data formats, and repurposing of data, all of which are central to GPO's Future Digital System (FDsys).

This Concept of Operations (Conops) document provides an overview of the proposed CSR. It is a living document that will evolve in collaboration with industry and GPO's stakeholders.

FINAL

## Table of Contents

1.	Scope .....	1
1.1	Identification .....	1
1.2	Document overview .....	1
1.3	System overview .....	2
2.	Referenced Documents .....	2
2.1	Standards and guidelines .....	2
2.2	GPO documentation .....	3
2.3	Laws and regulations.....	3
3.	Current Situation.....	3
3.1	Background, objectives, and scope of the current situation.....	3
3.2	Operational policies and constraints.....	5
3.3	Description of the current environment.....	6
3.3.1	Content ingest .....	6
3.3.2	Composition processes .....	8
3.3.3	Production/content delivery .....	9
3.3.4	Operational environment .....	10
3.3.5	Cost of system operations .....	10
3.4	Modes of operation.....	10
3.5	User classes and other involved personnel .....	10
3.5.1	Profiles of user classes.....	10
3.5.2	Interactions among user classes .....	11
3.5.3	Other involved personnel.....	11
3.6	Support environment .....	11
4.	Justification For and Nature of Changes.....	11
4.1	Justification for changes .....	12
4.2	Description of desired changes/capabilities.....	12
4.2.1	Proposed system attributes .....	12
4.2.2	Proposed system capabilities .....	13
4.2.3	Proposed system interfaces .....	13
4.3	Priorities among changes .....	14
4.4	Changes considered but not included .....	14
4.5	Assumptions and constraints.....	14
4.6	Adverse effects.....	14
5.	Concepts for the Proposed System .....	15
5.1	Background, objectives, and scope .....	15
5.1.1	Business objectives .....	15
5.1.2	Scope .....	16
5.2	Operational policies and constraints.....	16
5.3	Description of the proposed system .....	16
5.4	Modes of operation.....	17
5.5	User classes and other involved personnel .....	18
5.5.1	Capabilities.....	18

**FINAL**

5.5.2	Profiles of user classes.....	18
5.5.3	Administrative and support personnel .....	18
5.6	Support environment .....	19
5.6.1	System environment.....	19
6.	Operational Scenarios .....	20
7.	Summary of Impacts.....	21
7.1	Operational impacts.....	21
7.2	Organizational impacts .....	22
7.3	Impacts during development.....	22
8.	Analysis of the Proposed System .....	23
8.1	Summary of improvements.....	23
8.2	Disadvantages and limitations .....	23
8.3	Alternatives and trade-offs considered .....	24
	Appendix 1: Glossary.....	26
	Appendix 2: Acronyms .....	28

FINAL

## CONCEPT OF OPERATIONS (CONOPS)

### 1. SCOPE

This document describes the desired characteristics of Composition System Replacement (CSR) from the user's viewpoint. The sections below identify the proposed system, provide an overview of the document, describe the approach used to generate the document, and provide a brief overview of the system.

#### 1.1 IDENTIFICATION

The proposed CSR will encompass all of the associated equipment, facilities, material, software, hardware, policy and technical documentation, services, and personnel required for its operations and support.

#### 1.2 DOCUMENT OVERVIEW

The CSR Conops serves as a vehicle to communicate the high-level system characteristics of the envisioned system to users, developers, and other stakeholders. The CSR Conops was used to generate high level system requirements that are captured in the CSR Requirements Document (RD). The CSR Conops should be reviewed together with the CSR RD, as it contains additional information that has not been presented herein.

This document contains the following sections;

- **Section 1, Scope**, describes the approach used to develop this document.
- **Section 2, References**, lists the documentation that was used as a basis to create the document.
- **Section 3, Current GPO Situation**, describes existing GPO systems, operations, environment, and users.
- **Section 4, Justification For and Nature of Changes**, discusses the rationale for a Composition System Replacement and the nature of the proposed changes.
- **Section 5, Concepts for the Proposed System**, discusses design and operational concepts.
- **Section 6, Operational Scenarios**, describes examples of potential use of the system.
- **Section 7, Summary of Impacts**, describes operational, organizational, and other outcomes that could be expected to occur in the development of the system.
- **Section 8, Analysis of the Proposed System**, presents detailed system capabilities and operations.

FINAL

### **1.3 SYSTEM OVERVIEW**

Composition System Replacement (CSR) will be composed of the necessary technology and business practices to enable GPO to replace or integrate all existing discrete applications, utilities, and processes currently used by GPO and its users to compose and create files optimized for printing and access of select Congressional and Federal agency publications.

CSR will be aligned with GPO's *Strategic Vision for a 21st Century* and the agency's enterprise architecture (EA). It will also integrate with evolving enterprise-wide systems including GPO's Future Digital System (FDsys) (<http://www.gpo.gov/projects/fdsys.htm>), Oracle, and Digital Production System, in order to automate many content lifecycle processes, back office functions, and delivery of the content in formats suited to the needs of GPO customers.

CSR will, at a minimum, be capable of delivering files in XML, Postscript, Portable Document Format (PDF), and ASCII with locator codes to support GPO production and customer requirements. The system will also support essential features including line and page numbering, hyphenation, and justification. CSR must be capable of generating sophisticated tables giving the user extensive control over table formats, styles, etc.

CSR will permit data to be reused and repurposed. For example, a table tagged for a report should not have to be retagged in order to include it in the Congressional Record, a bill, or the U.S. Code.

The system must be capable of handling special characters including UTF-8 and entity codes. CSR applications and content must be accessible for all users including those with disabilities and follow best practices and regulations for accessibility (e.g. Section 508, W3C, etc.)

CSR will allow GPO and its users to employ XML while maintaining the current functionality of GPO locator codes. It is assumed that over time all users will migrate to the new XML workflow to take advantage of XML's flexibility and robustness.

## **2. REFERENCED DOCUMENTS**

Standards, guidelines, and other documentation used to support the *Composition System Replacement Conops* are described in sections 2.1 – 2.3.

### **2.1 STANDARDS AND GUIDELINES**

*IEEE guide for information technology-system definition-concept of operations (Conops) document. IEEE Std. 1362-1998* New York: Institute of Electrical and Electronics Engineers, 1998.

United States Government Accountability Office. *Government Printing Office: Actions to strengthen and sustain GPO's transformation.* GAO-04-830 Washington: U.S.

FINAL

General Accounting Office, 2004

World Wide Web Consortium (W3C). *Web content accessibility guidelines 1.0*  
Cambridge, MA [?]: W3C, 1999. <http://www.w3.org/TR/WCAG10/>

## **2.2 GPO DOCUMENTATION**

United States. Government Printing Office. Concept of Operations for the Future Digital System V2.0. 16 May 2005.  
[http://www.gpo.gov/projects/pdfs/FDsys\\_Conops\\_v2.0.pdf](http://www.gpo.gov/projects/pdfs/FDsys_Conops_v2.0.pdf).

United States. Government Printing Office. Requirements Document for the Future Digital System V2.1. 18 April 2006.  
[http://www.gpo.gov/projects/pdfs/FDsys\\_RD\\_v2.1.pdf](http://www.gpo.gov/projects/pdfs/FDsys_RD_v2.1.pdf)

United States. Government Printing Office. Government Printing Office Style Manual. 2000.

United States. Government Printing Office. A Strategic Vision for the 21<sup>st</sup> Century. Washington: U.S. Government Printing Office, 2004.  
<http://www.gpo.gov/congressional/pdfs/04strategicplan.pdf>

## **2.3 LAWS AND REGULATIONS**

"Production and Procurement of Printing and Binding" Title 44 *U.S. Code*, Chapter 5, 2000 edition

United States Cong. House. Making Appropriations for the Legislative Branch for the Fiscal Year Ending September 30, 2004, and for Other Purposes. 108th Cong., 1st sess., 2003. H. Rpt. 108-279. Washington: GPO, 2003.

## **3. CURRENT SITUATION**

This section describes the background, objectives, and scope of the current situation at GPO and the systems in use; operational policies or constraints; current environment; modes of operation; user classes and other involved personnel; and the support environment.

### **3.1 BACKGROUND, OBJECTIVES, AND SCOPE OF THE CURRENT SITUATION**

GPO's current composition system, Microcomp, was developed internally over 30 years ago to compose, set type and format pages of a large majority of Congressional products and select Federal agency publications. The system is a collection of 30 separate programs comprising a total of more than 212,000 lines of code. Microcomp's foundation is a GPO-developed data tagging scheme known as locator code.



**FINAL**

Over the last 10 years, GPO has continuously altered the original Microcomp applications to meet the evolving needs of its customers (e.g., processing of SGML/XML tagged data) and in-house print and electronic access requirements. In addition, approximately 700 related applications and utilities have been developed to enhance and sustain Microcomp composition processes and workflows.

GPO's ad-hoc implementation of new applications and utilities to support Microcomp has resulted in the development of disparate functionality and work processes. For example, GPO employees may use the latest version of the Microcomp application, GPO will also accept content from external customers using previous versions of Microcomp.

No continuity of operations plan (COOP) exists for Microcomp.

Both the Senate and the House of Representatives employ customized XML authoring tools based on XMetal software. The Senate version is a customized XMetal application known as Legislative Editing XML Application (LEXA). GPO has some development and support responsibilities for LEXA.

The following is a list of congressional products and agency publications composed with Microcomp:

- Acquisition Regulations
- Alphabetical List of Senators
- Appropriations Budget Estimates
- Architect of the Capitol
- Attorneys and Agents Registered to Practice Before the U.S. Patent and Trademark Office
- Biographical Directory of the U.S. Congress
- Briefs - Commission on Security and Cooperation in Europe
- Catalog of Federal Domestic Assistance
- Code of Federal Regulations
- Committee hearings
- Committee prints
- Committee print compilations
- Congressional bills and resolutions
- Congressional Documents
- Congressional Directory
- Congressional Record
- Congressional Record Index
- Congressional Reports
- House Resolutions
- House Rules and Manual
- How Are Laws Are Made
- Index Medicus
- Interior Decisions
- Justice Department - Administrative Decisions
- Library of Congress Telephone Directory
- Medicare Commission Report
- Our American Government
- PEPS (In-House)
- Production Office
- Patents and Trademarks
- Plum Book
- Privacy Act (OFR)
- Proclamation - U.S. Statutes at Large
- Public and Private Laws
- Records Management Handbook for U.S. Senators and their Archival Repositories
- Riddick's Senate Procedure

**FINAL**

- Congressional Research Service
- Constitution of the U.S.A.
- Defense Dept. Telephone Book
- Deshler's Precedents
- DOD Logistics
- Economic Indicators
- Economic Report of the President
- Eulogies
- Federal Register
- Federal Travel Regulations
- GPO Style Manual
- Hearings/Reports
- History of Senate Finance Committee
- History of House Committee on Appropriations
- History of the House of Representatives
- House and Senate Calendars
- House and Senate Telephone Directories
- House Commerce Committee
- House Journal
- House Parliamentarian
- House Practice
- House Printing Services
- Rules Booklets
- Senate Agriculture Committee
- Senate Banking Committee
- Senate Committee and Subcommittee Assignments
- Senate Ethics Manual
- Senate Foreign Relations Anniversary
- Senate Legislative Calendars
- Senate Intern Handbook
- Senate Journal
- Senate Manual
- Senate Telephone Book
- Social Security Acquiescence Ruling
- Supreme Court Briefs
- Thumb Index for Committee Prints
- Tributes
- Unified Agenda
- Unveilings
- U.S. Budget
- U.S. Code
- U.S. Government Manual
- U.S. Tax Court Monthly Reports
- Weekly Compilation of Presidential Documents

### **3.2 OPERATIONAL POLICIES AND CONSTRAINTS**

Microcomp functions in an environment of ever-growing limitations and constraints, coupled with a fundamental aversion to change from both internal and external stakeholders, since Microcomp has met these users' needs for several years.

From the GPO perspective, Microcomp's locator tagging scheme requires GPO staff to maintain expertise about a technology that is only supported internally and is not based upon industry best practices. The current system is in operation 24/7 and therefore requires 24/7 support. GPO has lost several Microcomp subject matter experts and there are no immediate plans to train additional support staff. GPO's Information Technology & Systems (IT&S) group no longer enhances or upgrades Microcomp, although basic software support is still provided.

Because several versions of Microcomp are currently used by external customers, GPO does not have a complete understanding of what operating systems, hardware, other applications, customizations, and plug-ins the current system interacts with.

Finally, the current system does not support GPO customer's need to publish documents directly from SGML/XML.

### 3.3 DESCRIPTION OF THE CURRENT ENVIRONMENT

The detailed current-state documentation is based on a simple reference model, depicted in Figure 1, which describes the content workflow through the current composition system.

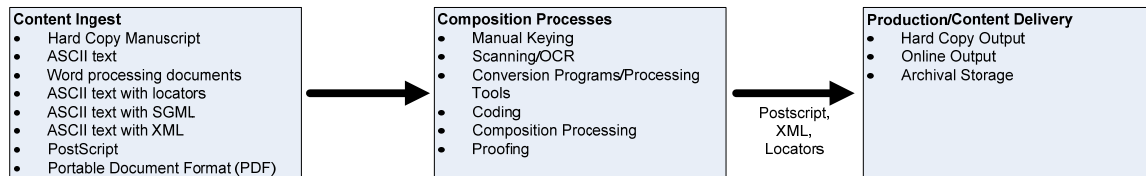


Figure 1. Simple Reference Model

#### 3.3.1 CONTENT INGEST

Data files are accepted from Content Originators in a variety of formats for composition of publications. Content is provided in the following formats:

- Hardcopy manuscript
- ASCII text
- Word processing documents
- ASCII text with locators
- ASCII text with SGML
- ASCII text with XML
- PostScript
- Portable Document Format (PDF)

#### Typical Workflow for Hard Copy Manuscript

1. Hard copy is mailed or delivered to the Congressional Publishing Services at GPO. The hardcopy manuscript is the official (master) version of the content, regardless of format received.
2. Hard copy is forwarded to the Mark up Section where a determination is made whether the publication will be optical character recognition (OCR) scanned or rekeyed.
  - SCANNED: OCR is applied to the scanned files. The Mark-up Section handwrites on the hard copy where locator or XML tags should be added to the file, and divides the hard copy manuscript into segments, to allow multiple operators to work on the job at the same time. The Keyboard Section divides the file into multiple files matching the hard copy manuscript segments. Multiple keyboard operators are assigned the manuscript segments and corresponding files to insert the tags using Xywrite or Textpad.
  - REKEY: Documents are sent to the Mark up Section. The Mark up Section handwrites on the hard copy where locator or XML tags should

**FINAL**

be added to the file and divides the hard copy manuscript into segments to allow multiple operators to work on the job at the same time. Multiple keyboard operators are given the manuscript segments and they rekey the document and add tags using Xywrite or Textpad. Multiple files are generated matching the hard copy manuscript segments.

*NOTE: When XML tags are present, they are converted to locators during the workflow, depending upon the type of publication. For example, Congressional Bills are edited as XML and converted to locators for composition; the Code of Federal Regulations is edited and composed as SGML and converted to locators at the end of the process to create access copies for WAIS.*

3. Tagged locator files are saved to a hot folder and automatically composed as galley proofs using Microcomp.
4. PostScript is output, and the galley proofs are printed from the PostScript directly to the Proofroom printer.
5. Proofreaders compare the galley proofs to the hard copy manuscript. Corrections are written on the galley proofs.
6. Galley proofs are reviewed by the referee who answers any questions written on the galleys, gathers all galleys and hard copy manuscript from a single job, ensures the style of corrections is consistent, and sends the entire compiled group of galley proofs and manuscript back to the Keyboard Section.
7. The Keyboard Section makes corrections to the content and original tags (XML or locators).
8. All file segments for a job are combined into a single file (SGML / XML); Locator files do not have to be combined.
9. The corrected locator file is composed as pages using Microcomp (see Section 3.3.2), PostScript is output, and page proofs are printed from the PostScript. Individual PostScript files are created for each page of a publication.
10. Page proofs are sent to the Proofroom. The galley proofs are compared to the page proofs.
11. CORRECTIONS NEEDED: Marked up page proofs are sent back to Keyboard for correction. Repeat output of page proofs and file corrections until approved. Individual pages can be recomposed without recomposing the entire publication.
12. PostScript files for individual pages are combined into a single, complete publication.
13. PostScript file is sent direct-to-plate or to the Digital Print Center. The Digital Print Center may distill to PDF if necessary.
14. PostScript file is distilled to PDF by the Text Processing Section for Internet distribution.
15. Internet access files (ACSII and PDF) are created from the corrected (final, approved) locator file by the Text Processing Section.
16. Corrected locator or XML file and the PostScript file are archived.

**Typical Workflow for Untagged ASCII Text or with Locator Tags, SGML Tags or XML Tags**

1. Files are sent to the Markup Section via a webpage or file transfer (e.g., e-mail, FTP), or to the Congressional Office via digital media. Numerous FTP sites and

**FINAL**

- customer accounts exist. Corresponding hard copy manuscript is mailed or delivered to the Congressional Office at GPO.
2. The ASCII files and the hard copy manuscript are sent to the Mark-up Section. The Mark-up Section handwrites on the hard copy where locator or XML tags should be added to the file, and divides the hard copy manuscript into segments to allow multiple operators to work on the job at the same time. The Keyboard Section divides the file into multiple files matching the hard copy manuscript segments. Multiple keyboard operators are given the manuscript segments and corresponding files and they insert the tags using Xywrite or Textpad.

*NOTE: When XML tags are present, they are converted to locators at some point (or at multiple points) during the workflow process. This is dependent upon the type of publication (e.g., Congressional Bills are edited as XML and converted to locators for composition; the Code of Federal Regulations is edited and composed as SGML and converted to locators at the end of the process to create access copies for WAIS.)*

3. Follow steps 3-15 from Typical Workflow for Hard Copy Manuscript.

**Typical Workflow for Word Processing Documents**

1. Convert Word Processing Document to ASCII and follow Typical Workflow for Untagged ASCII file.

**Typical Workflow for PostScript and Portable Document Format (PDF)**

1. These files are already OK to print and sent directly to Digital Print Center or Direct to Plate. These are files that are not a piece of a larger publication.
2. These files are usually accompanied by a locator database that is processed for posting on the Internet.

GPO accepts various file formats through a variety of methods including hard copy manuscript, electronic transfer (e.g. File Transfer Protocol (FTP) or on digital media (e.g. compact disc (CD)). Content processing depends on the file format and medium utilized.

**3.3.2 COMPOSITION PROCESSES**

Composition processes apply specific publication styles and settings to ingested content for printing and screen presentations:

1. Text Processing Section ingests tagged ASCII, SGML, and XML into Microcomp.
2. Microcomp composes a data file by applying a specific composition format to tagged data. The format file contains all the typographic attributes and information required to produce a fully composed page(s) in PostScript. This information includes: typeface, page layout, values of the locators, automatic footnote numbering, auto-generated running headers, and page numbering.
3. Graphic Systems Development Division (GSDD) and Text Processing Section apply Microcomp utility processes to tagged files and Postscript for various dissemination purposes (e.g., to get file to print properly, GPO Access WAIS).

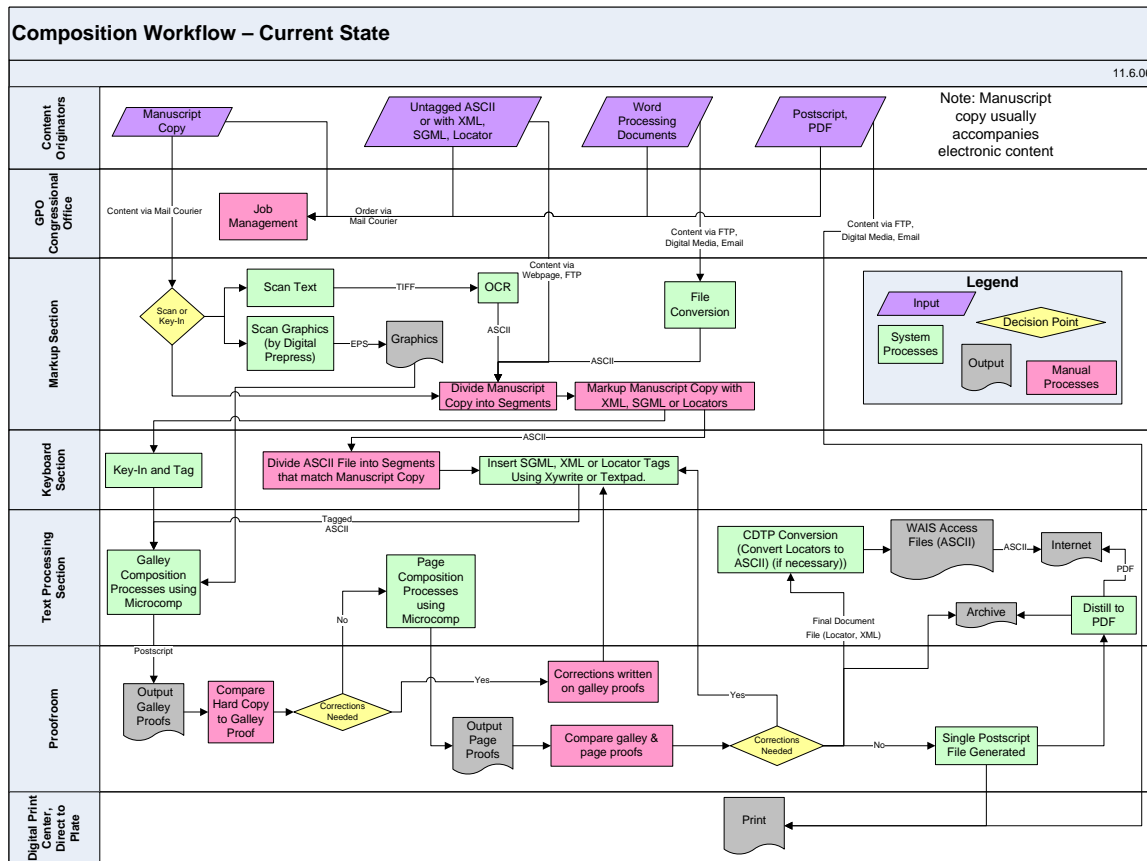
FINAL

### 3.3.3 PRODUCTION/CONTENT DELIVERY

Composed content is rendered for multiple forms of delivery:

1. Once approved, the final proofed copy/coded data (in Postscript) is used to produce offset plates for traditional print production
2. The Postscript is also distilled into PDF for publication to the Internet. Online publishing of various products is accomplished through batch processing.
3. The locator file is re-processed using CD-Typesetting Program (CDTP) to reflow the text and generate ASCII files for placement online. The coding is logically structured so that it approximates the look of the printed publication within the limitations of ASCII.
4. Locator files and PDF files are archived by GPO.
5. Locator files for certain products (e.g., Bills, Congressional Record, Congressional Record Index) are sent via FTP to the Library of Congress and other external users. Other files are available to external content originators for repurposing.

Figure 2 below depicts the current composition workflow. The current environment is closely tied with Plant Operations systems and operational areas, including Markup Section, Keyboard Section, and others. The process is largely manual, and workflows vary by type of publication.





FINAL

## **Figure 2. Composition Workflow – Current State**

### **3.3.4 OPERATIONAL ENVIRONMENT**

The current operating environment consists of a server cluster for document and graphic file storage and several networked PCs. The servers are located in the Plant Operations organization in Washington, DC. Approximately 150 users utilize 200 Windows 2000 workstations located throughout Pre-Press Division. In addition, approximately 100 customers elsewhere in Federal government use Microcomp on a daily basis on their own computers. File backups, documentation, graphics, and applications for in-house work are maintained by GPO personnel.

### **3.3.5 COST OF SYSTEM OPERATIONS**

Information regarding the cost of current system operations is not available.

## **3.4 MODES OF OPERATION**

Current practices provide limited automation for processing and preserving content, requiring human intervention at many steps in the process. In the normal operational environment, all production using the Microcomp system is at the main GPO facility in Washington, D.C. GPO has an alternate site for limited production and composition using Microcomp located within 50 miles of Washington, D.C. Similarly, GPO utilizes another alternate site solely for data backup purposes located within the same 50 mile radius.

## **3.5 USER CLASSES AND OTHER INVOLVED PERSONNEL**

### **3.5.1 PROFILES OF USER CLASSES**

- Content Originators (e.g., House of Representatives, Senate, Office of the Federal Register, Congressional Composition Operators): Develops and gathers information and content and generates requests for GPO services.
- GPO Congressional Publishing Services: Accepts hard copy manuscript, electronic files on CD, DVD, floppy disk, assigns jacket numbers, supply specifications and distributions as necessary, forwards work to Plant Operations, resolves any problems or queries with work during production and delivery.
  - User Support Team: Installs Microcomp and provides software support.
- GPO Plant Operations
  - GPO Prepress Division
    - Markup Section: Provides scanning and file conversion services and prepares manuscript copy for Keyboard Section.
    - Keyboard Section: Keys and tags content using Xywrite or TextPad.
    - Proofroom: Provides proofing and editing services
    - Digital Print Center, Direct to Plate: Provides printing services

## FINAL

- Text Processing Section: Applies composition specific format to content using Microcomp and is responsible for Microcomp product support.
- GPO Graphic Systems Development Division (GSDD): Provides in-house user support, develops and maintains composition specific formats, typesets selected work.

### 3.5.2 INTERACTIONS AMONG USER CLASSES

The interactions of user classes the system is highlighted in Figure 2 of Section 3.3.

### 3.5.3 OTHER INVOLVED PERSONNEL

- Information Technology and Systems (IT&S)
  - Infrastructure: Maintains and administers servers and network connections.
  - User Support: Provides user support/help desk operations.
  - Configuration/Change Management: Responsible for change/configuration management of Microcomp.
  - System Development: Program maintenance for Microcomp and LEXA Authoring Tool.
  - Product Support: XML Bill DTD development and maintenance, GPO support for LEXA

## 3.6 SUPPORT ENVIRONMENT

The existing composition system is supported by in-house developers. Each system has its own unique support environment. Current systems do not meet GPO needs and it would be cost prohibitive to upgrade or modify; thus the following information is omitted from this document:

- Identification of the support concepts
- Identification of the support environment for the current system, including:
  - Support agency or agencies
  - Facilities
  - Equipment
  - Support software
  - Repair or replacement criteria
  - Maintenance levels and cycles
  - Storage
  - Distribution
  - Supply Levels

## 4. JUSTIFICATION FOR AND NATURE OF CHANGES

The following subsections present justification for changes, the nature of the desired changes, priorities among changes, changes considered but not included, and



**FINAL**

assumptions and constraints associated with building the system.

## **4.1 JUSTIFICATION FOR CHANGES**

Many factors drive the desire to replace the current composition system, including the customer's need for GPO to accept, compose, and output XML; advances in technology to improve the workflow processes of content originators and composition operators; and the need to meet requirements for ingest into FDsys.

The current process for accepting XML and converting it to GPO locators is inefficient and adds additional potential points of failure. A system that can accept SGML, XML, and locators, compose content, and output XML will provide GPO with the means to meet the needs of Congressional and Federal agency customers and end users.

CSR will allow GPO users the freedom to migrate to XML incrementally, with the eventual elimination of locators (and subsequently Microcomp). The transition to XML will enable the users to take advantage of all the features of XML including enhanced search and retrieval capabilities and repurposing of data.

Major arguments for change are:

- GPO needs to reduce its dependency on an internally-developed method of composition that is not based on an industry standard data model, has no IT support outside of GPO, and has limited IT support inside GPO.
- The majority of external users are dissatisfied with the practice of converting XML back to locators to typeset, therefore not utilizing XML to its full capacity.
- There are multiple points of failure in the current system with few alternatives for repair or work-around.

## **4.2 DESCRIPTION OF DESIRED CHANGES/CAPABILITIES**

The system will accept: XML and locator code input, thereby eliminating the need for separate conversion programs for each publication database.

The system will output: XML for optimal search/retrieval/repurposing in FDsys, Postscript/PDF/HTML to support customer printing/electronic display requirements, and locators.

The system will align with GPO's Enterprise Architecture, and interface with FDsys, and existing pre- and post-processing applications. The system will be based on an XML data model to enable support, maintenance, and development outside of GPO.

### **4.2.1 PROPOSED SYSTEM ATTRIBUTES**

The system should possess the following attributes:

## FINAL

- **Modularity.** Ability to use plug-in components that can be replaced with minimal impact to remaining components as workload and technology change;
- **Scalability.** Capable of accommodating increasing volumes of content;
- **Extensibility.** Ability to handle additional file formats over time, not limited to specific types that exist today;
- **Comprehensiveness.** Accept, process, and output a variety of file formats as outlined in this Conops;
- **Flexibility.** Enable GPO to tailor content-based services to suit its customers' needs and enable GPO to implement progressive improvements in its business process over time; and
- **Accessibility.** Employ software and create output compliant with Section 508 of the Rehabilitation Act.

#### 4.2.2 PROPOSED SYSTEM CAPABILITIES

To meet strategic objectives, the new system will need to:

- Accept XML and locator code input.
- Output XML, Postscript, PDF, HTML, locators.
- Support existing and new style formats.
- Support long-term phase-out of locators for input and output.
- Automate existing workflows and processes.
- Track jobs throughout the composition process.
- Support batch processing of content.
- Support system security.
- Manage upgrades to software.
- Provide audit trails and logs of system activity.
- Align with agency IT strategy and GPO's Enterprise Architecture (EA).

#### 4.2.3 PROPOSED SYSTEM INTERFACES

The system must interface with the following external applications and systems:

- Content authoring tools (customer remote locations and GPO)
  - LEXA
  - TextPad
  - XMetaL
  - Xywrite
- GPO's Future Digital System (FDsys)

**FINAL**

Internal applications and systems that the composition system must interface with will be defined in the *CSR RD*.

### **4.3 PRIORITIES AMONG CHANGES**

The solution should be able to accept both XML and locator files as input, and output XML, locators, Postscript, and PDF while retaining the existing style formats used by GPO's customers.

### **4.4 CHANGES CONSIDERED BUT NOT INCLUDED**

See Section 8.3 Alternatives and Tradeoffs for a summary of alternatives to Composition System Replacement and changes considered.

### **4.5 ASSUMPTIONS AND CONSTRAINTS**

1. The CSR data tagging model will be SGML/XML based.
2. The CSR will interface with the Future Digital System (FDsys).
3. The system will address usability and accessibility requirements and best practices.
4. The system will be flexible, extensible, and adaptable.
5. Tools and processes may include human interactions into the foreseeable future.
6. Customers will have confidence in GPO's management and processing of their content.
7. CSR implementation will require extensive training of both internal and external users.

### **4.6 ADVERSE EFFECTS**

The risks of not proceeding with CSR include:

- Inability to fulfill agency mission and meet customer needs
- Inability to implement industry technology advances and best practices
- Eventual inability to repair or develop workarounds for production workflows
- Unstable processing of standard formats (e.g., XML)
- Inability to integrate Microcomp with current GPO strategic initiatives (e.g., FDsys)

**FINAL**

- Loss of institutional support and programming resources specific to Microcomp
- GPO will not be positioned to provide adequate guidance, assistance, or services to customers

## **5. CONCEPTS FOR THE PROPOSED SYSTEM**

The following subsections describe the proposed applications with respect to background, objectives, and scope; describe applicable operational policies and constraints; describe the proposed system; describe user classes and other involved personnel; and describe the support environment.

### **5.1 BACKGROUND, OBJECTIVES, AND SCOPE**

A high level system overview has been provided in Section 1.3, System Overview. Goals and motivation for the new system are discussed in Section 3.1.2, Motivation for a New System.

#### **5.1.1 BUSINESS OBJECTIVES**

The objectives of the new system are:

- To implement a composition system that utilizes best practices and industry standard data models (e.g., XML)
- To continue to support traditional GPO and customer composition input formats (e.g., locators) until best practices and industry standards are fully adopted
- To support multiple output formats (e.g., PDF, PostScript, XML, locators) as required by GPO and customers
- To support a flexible, consolidated, extensible, scalable, and efficient Production workflow
- To allow necessary updates or to incorporate changes to workflow processes on an as-needed basis as they are developed and implemented
- To achieve alignment with GPO IT strategy and enterprise-wide endeavors such as FDsys, GPO Enterprise Architecture, the GPO Oracle Project, and Digital Production Systems
- To support the capability for content to be reprocessed in its original format or repurposed as required
- To provide access to CSR applications and resulting publications to all users, including those with disabilities
- To provide training for GPO and agency users on the new composition system
- To minimize disruption of customer and GPO processes

FINAL

### **5.1.2 SCOPE**

This application is expected to replace GPO's Microcomp legacy system to meet the evolving needs of GPO's customers and changing technology. CSR will accept content tagged with SGML/XML and will accept ASCII tagged with GPO's legacy locator codes.

The new application(s) will replace all of GPO's current legacy composition applications and their approximately 700 associated support utilities, to comprise a complete, end-to-end composition application including authoring tools and processes, pre-composition processing, composition, WYSIWYG proofing, post-composition processing, and output. CSR will integrate with GPO's FDsys for primary functions including content ingest, content management, workflow, content delivery, and access.

## **5.2 OPERATIONAL POLICIES AND CONSTRAINTS**

The proposed composition system will support GPO operational policies as well as future changes to policy.

Constraints that may have an impact on system design are:

- Composition system design must interface with or support FDsys
- System must be designed and implemented in accordance with the strategic technology and architecture decisions of the GPO Enterprise Architecture
- Implementation of CSR will need to be consistent with the GPO Information Assurance Security Policy, with special consideration given to the operational access to system data to various GPO personnel
- The system will need to integrate with current customer authoring and composition workflow processes (e.g., LEXA) to allow minimal disruption
- The system will be implemented within current funding and resource allocations
- CSR will adhere to all applicable laws and regulations (e.g., Section 508)
- Users of the proposed system will need to be trained
- CSR will need to be operated both as a composition server and as a stand-alone desktop application

## **5.3 DESCRIPTION OF THE PROPOSED SYSTEM**

CSR is a component of the Pre-Ingest Processing component of FDsys. Input to the system will come from many sources including both internal and external users. The composition engine will process input data by applying specified styles and will be capable of generating both hard and soft-copy proofs.

Upon approval of proofs, the composition engine will output content in various formats which will be sent to GPO's Plant Operations for hard-copy production.

FINAL

The composition engine will output PostScript, XML, and PDF for FDsys processing. CSR will apply necessary processes to make files accessible in HTML and PDF for GPO Access, subscribers, and other data feeds.

System metadata about the composition process will be recorded for use in FDsys.

CSR will be a component of FDsys Pre-ingest Processing which manages the functions that prepare content for ingest into the system (figure 3).

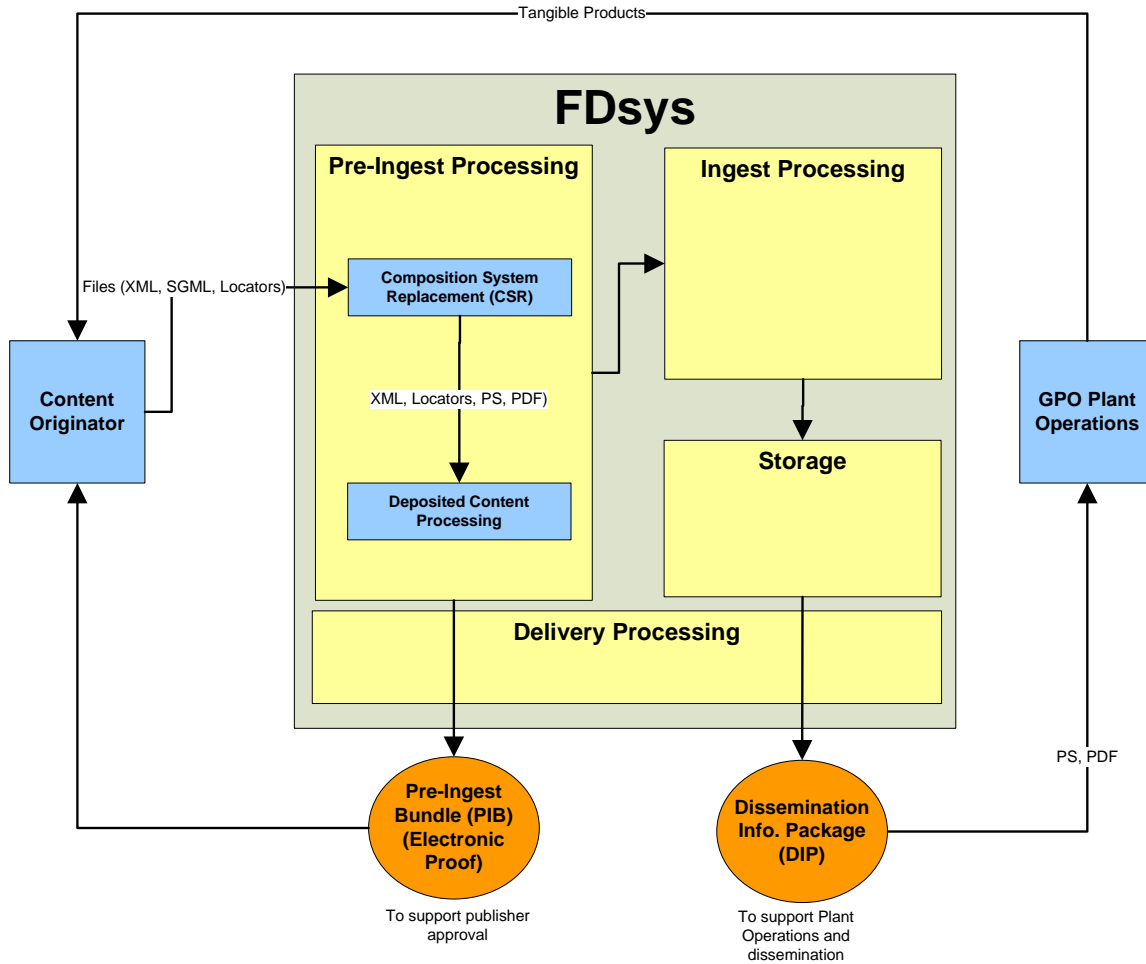


Figure 3: CSR & FDsys Functional Model

## 5.4 MODES OF OPERATION

The modes of operation for the proposed system as currently known are:

**Nominal.** Production of all products using the composition replacement system would be at agency customer sites and at the main GPO facilities in Washington, DC. The system would use both automated and manually initiated processes.

**FINAL**

**Maintenance.** Testing of system component upgrades, replacements, etc. would occur in a separate development test environment available to users as needed and would have roll-back capabilities. Specific user classes will have the option to accept or decline upgrades to the system. Versioning will remain consistent between stand-alone and server implementations of the composition engine.

**Alternate site.** Implemented when one site has a failure. A disaster recovery alternate site should be operational with minimal delay following a primary site failure. This site would be in compliance with GPO's Continuity of Operations Plan.

## **5.5 USER CLASSES AND OTHER INVOLVED PERSONNEL**

The following subsections describe the organizational structure and the classes of users, including user capabilities, which are associated with the proposed system.

### **5.5.1 CAPABILITIES**

The capabilities for the primary function within the user classes of the system are described above in the subsection of section 5.3, Description of Proposed System.

### **5.5.2 PROFILES OF USER CLASSES**

Content Originators (Congressional & Agency customers) will:

- Create content
- Submit content, or make content available, to GPO

Composition Operators (Congressional & Agency customers, Pre-Press Divisions: Keyboard, Markup, Proofroom) will perform:

- Keyboard operations, including the initial entry of content and tagging as well as corrections
- Mark-up and copy preparation operations
- OCR scanning of manuscript
- Scanning of graphics
- File translations
- Soft and hard copy proofing
- Composition generation of both galley and page output

### **5.5.3 ADMINISTRATIVE AND SUPPORT PERSONNEL**

Formatting and Style Sheet Developers (GSDD) will perform:

- DTD and/or Schema design definition
- Style sheet and format development (publication specific setting and processes)
- XSLT development
- Pi font development

**FINAL**

Composition System Administrators (GSDD, Text Processing) will:

- Develop publication specific automated workflows including both pre and post composition processes
- Administer the batch operation of composition servers, automated processes and print queues
- Manage rollout/rollback of new system components
- Composition Application Support (first level)

Information Technology & Systems (IT&S) will perform:

- User Support/Help desk operations
- Change/configuration management
- Maintain and administer servers and network connections

## **5.6 SUPPORT ENVIRONMENT**

The support environment will not be determined until completion of the system analysis and design phase of the CSR program. The CSR Conops will be updated as required.

The Composition System Replacement will be supported at the same level as the existing composition system is supported (e.g., 24 x 7) for all users.

### **5.6.1 SYSTEM ENVIRONMENT**

CSR must be capable of residing on current GPO systems and future EA compliant systems.

- Hardware Platforms
  - UNIX
  - Windows servers
  - PC Desktop
- Operating Systems
  - UNIX
  - Windows
- Web and Application Server (if necessary)
  - UNIX
  - Windows servers



FINAL

## **6. OPERATIONAL SCENARIOS**

### **CONTENT ORIGINATOR SCENARIOS**

#### **Scenario: External Content Originators with XML Authoring Tool**

- Create and tag content using text editors or authoring tools that generate XML
- Use CSR to generate proofs
- Make corrections and modifications to content based on proofs
- Approve content for printing
- Submit content to GPO
- Receive or retrieve files from GPO after print production
- Use content received from GPO to create new versions of content

#### **Scenario: External Content Originators with Word Processing Software**

- Create content using word processing software
- Use word processing program to generate proofs
- Make corrections and modifications to content based on proofs
- Approve content for printing
- Submit content to GPO

#### **Scenario: External Content Originators using Xywrite or Textpad**

- Create and tag content with GPO locators using Xywrite or Textpad
- Use CSR to generate proofs
- Make corrections and/or modifications to content based on proofs
- Approve content for printing
- Submit content to GPO
- Receive or retrieves files from GPO after print production
- Use content received from GPO to create new versions of content

#### **Scenario: External Content Originators Submitting Manuscript**

- Gather manuscript from various sources where electronic files are not available
- Submit manuscript or camera copy to GPO

#### **Scenario: External Content Originators Notify GPO of Content on Web site**

- Create and tag content as HTML
- Notify GPO of location of content on Web site
- Receive and approve proofs for publication

Note: an individual job may involve multiple Content Originator Scenarios listed above.

### **COMPOSITION OPERATOR SCENARIOS**

#### **Scenario: Composition Operators**

- Receive content from Content Originator(s)
- Identify publication type (e.g., Congressional Bill) and publication (e.g., HR 3618)

**FINAL**

- Identify current tagging format (e.g., ASCII, Word Processor)
- Identify desired tagging format (e.g., XML)
- Translate tagging format if required (e.g., locators or SGML into XML)
- Manually apply additional tagging as needed
- Use CSR to generate proofs in galley and page format
- Review and edit formatting
- CSR generates and delivers final versions in PostScript, XML, and native document format

*Process Variations and Exception Processes for Manuscript:*

- Markup of manuscript with tag locations
- Keyboard operations, including the initial entry of content and tagging as well as corrections
- OCR scanning and automated or manual tagging of manuscript

**Scenario: Formatting and Style Sheet Developers**

- Develop or modify DTDs and Schema for use by CSR and authoring tools
- Use CSR to develop or modify publication specific settings and processes for use by CSR and authoring tools
- Develop or modify XSLTs for translating from one version of XML to another
- Use CSR to develop or modify XSL style sheets to render XML files within a browser
- Develop or modify Pi fonts for special characters

**Scenario: Composition System Administrators**

- Use CSR workflow tools to develop and modify publication specific automated workflows including both pre- and post-composition processes
- Administer the batch operation of composition servers, automated processes, and print queues
- Manage rollout/rollback of new system components
- Composition application support

**Scenario: Information Technology & Systems (IT&S)**

- Perform user support/help desk operations
- Change and configuration management
- Maintain and administer servers, network connections, and hardware

## **7. SUMMARY OF IMPACTS**

Implementation of CSR will have wide ranging impact on both GPO and its customers. The subsections below identify potential operational, organizational, and development impacts that should be considered as GPO develops its plans.

### **7.1 OPERATIONAL IMPACTS**

Until CSR undergoes systems analysis and design, operational impacts of the proposed

**FINAL**

system are not known. Possible operational impacts include:

- New interfaces with primary or alternate computer operating centers
- Interfaces and interoperability with legacy systems and applications
- Changes in processes and procedures
- Use of new data types (e.g., XML)
- New data sources (e.g., customers, publications)
- New modes of operation during emergency, disaster, or accident conditions
- New methods (e.g., authoring tools) for creating and delivering input data
- Changes in operational budget
- Changes in operational risks

It is anticipated that GPO will have to implement changes to the way it conducts business in order to achieve the agency's mission, goals, and objectives. The system will facilitate this endeavor.

## **7.2 ORGANIZATIONAL IMPACTS**

GPO may have to develop or modify policies and business practices. The depth and breadth of the impact on number of personnel, skill levels, position identifiers, and locations of personnel is unknown at this time. The interaction of personnel with the system may require revision of position descriptions to reflect changes in GPO's business practices.

With this in mind, GPO has identified a number of possible organizational impacts:

- The commitment of resources (e.g., funding, time, staff) by GPO to establish working relationships with Congress and other Government agencies as the system is rolled out
- The development of education and increased training for both GPO staff and End Users
- System operation and maintenance
- Increased opportunities for career development for GPO staff

## **7.3 IMPACTS DURING DEVELOPMENT**

The full extent of development impact will be discovered in the systems analysis and design phase. Factors considered thus far include:

- Articulation of business rules and other controls needed for operational implementation
- Development of training plan to be implemented incrementally
- Training necessary for each release
- Involvement in studies, meetings, and discussions prior to award of the contract
- User Support and involvement in reviews and demonstrations, evaluation of initial operating capabilities and evolving versions of the system, development or

## FINAL

- modification of databases, and required training
- Parallel operation of test systems without disruption of current processes
- Parallel operation of the new and existing systems
- Operational impacts during system testing of the proposed system

## 8. ANALYSIS OF THE PROPOSED SYSTEM

This section describes the improvements, disadvantages and limitations, and alternatives and tradeoffs considered.

### 8.1 SUMMARY OF IMPROVEMENTS

#### **Proposed Approach: Native Composition of XML and Locators**

Under this approach the native file used for composition will be ASCII text tagged with SGML, XML, or locators. Input files are converted to and output as XML, Postscript, and PDF. In cases where GPO locator files are used as the native document format, CSR will deliver the locator files as part of the composition output process.

This solution was chosen for the following reasons:

- The result will be structured XML that can be ingested into FDsys, which will enable its users to take advantage of all the features of XML including enhanced search and retrieval capabilities and repurposing of data
- Microcomp will be phased out when CSR is being used to compose *all* products previously composed using Microcomp

Improvements to the current workflow process include:

- Consolidation of the number and types of conversions / scripts employed in Production today
- Consolidated, automated, and documented workflow processes
- Accurate reporting and accountability resulting from improved tracking and recording mechanisms
- Greater support is available for COTS applications
- Improved systems maintenance resulting from using standard EA sanctioned equipment
- Use of hardware and software platforms consistent with the GPO Enterprise Architecture

Summary of the proposed approach:

- Composes XML and locator files: Yes
- Outputs XML: Yes
- After CSR is fully operational, Microcomp can be retired: Yes

### 8.2 DISADVANTAGES AND LIMITATIONS

Potential disadvantages and limitations to this approach include:

## FINAL

- It may not be feasible or cost effective to develop a CSR that will use ASCII files tagged with either locators or XML as the native file format
- The formatting and layout from CSR may not exactly match existing output. This is unacceptable to GPO's Congressional customers
- The conversion of locators to XML may result in XML that is not as rich as content originally created using an XML authoring tool. This may not meet search and repurposing requirements for FDsys

### 8.3 ALTERNATIVES AND TRADE-OFFS CONSIDERED

Three alternatives have been outlined in the event that the selected approach is deemed unsuitable based on potential disadvantages or limitations of the selected approach.

#### Alternative 1: Translation of Locators to XML before Composition

Native files used for composition will be ASCII text tagged with XML. Locator files will be converted to XML using a conversion program external to CSR prior to the composition process. The composition process will only accept XML files, including the XML files converted from locators by the conversion program. Output will be XML, Locators, Postscript, and PDF.

Advantages to Alternative 1 include:

- Structured XML for ingest into FDsys, similar to the proposed approach
- Not requiring the composition process to accept and convert locators to XML should increase the number of vendor solutions and reduce the cost of the system

Potential disadvantages and limitations to Alternative 1 include:

- The formatting and layout from CSR may not exactly match existing output. This is unacceptable to GPO's Congressional customers
- The conversion of locators to XML using a conversion program may result in XML that is not as rich as content originally created using an XML authoring tool. This may not meet search and repurposing requirements for FDsys

Summary of Alternative 1:

- Composes XML and locator files: No- locator files are converted to XML prior to composition
- Outputs XML: Yes
- After CSR is fully operational, Microcomp can be retired: Yes

#### Alternative 2: Parallel Workflow

CSR will only accept XML files, not locator files. There will be no conversion of locator files to XML. Locator files will be composed using Microcomp, which will run in parallel with the CSR.

Advantages to Alternative 2 include:

- Not requiring the CSR to accept and convert locators to XML should increase the number of possible solutions and reduce the cost of the system

**FINAL**

- Not converting locator files to XML may encourage Content Originators to migrate their databases to XML at an accelerated pace

Potential disadvantages and limitations to Alternative 2 include:

- Microcomp must remain operational as long as databases with locators exist
- Current Microcomp skills and knowledge are diminishing within GPO
- XML will be incomplete, so users of FDsys will not be able to take full advantage of the features offered by XML (e.g., enhanced search and retrieval capabilities, repurposing of data)
- Content Originators will have to use two different programs to compose XML and locator files
- The formatting and layout from CSR may not exactly match existing output. This is unacceptable to GPO's Congressional customers
- The PDF output from locator files composed with Microcomp may not match the PDF output from XML files composed with the CSR. The PDF output from CSR will provide a number of features (e.g., bookmarks, hypertext links) that the PDF output created from Microcomp will not have unless manually created
- Content Originators that rely on legacy locator databases will be faced with the task of migrating their databases to XML

Summary of Alternative 2:

- Composes XML and locator files: No- cannot compose locator files
- Outputs XML: Only for XML files. Locator files will be not be converted to XML.
- After CSR is fully operational, Microcomp can be retired: No- not until all databases have been migrated to XML.

**Alternative 3: Typesetting DTD**

XML and locator files are translated by CSR to a generic DTD for typesetting and output as PDF/PostScript.

Advantages to Alternative 3 include:

- Composes both XML and locator files
- Open and flexible

Potential disadvantages and limitations to Alternative 3 include:

- The formatting and layout from CSR may not exactly match existing output. This is unacceptable to GPO's Congressional customers
- Mixed databases of XML and locator files will have limited search, retrieval, and repurposing capabilities
- Generic DTD does not contain the required fields

Summary of Alternative 3:

- Composes XML and locator files: Yes
- Outputs XML: No
- After CSR is fully operational, Microcomp can be retired: Yes

FINAL

## APPENDIX 1: GLOSSARY

**Batch Processing:** Automated page layout of a designated group of files based on pre-determined style rules.

**Entity Codes:** SGML/XML codes for representing special characters in ASCII.

**Frame:** GPO term for a composition page unit. The number of frames in a publication differs from the number of pages by the number of blank pages. Frame numbers always are numbered consecutively starting from '1' for the first printed page in the publication. Page numbers may or may not start at the first printed page (i.e. because preliminary pages are often numbered i, ii, iii, etc.), or only restart at '1' in the first issue of the year (e.g., for the Federal Register).

**Frame Number:** The sequential number identifying a composition page unit.

**Folio:** A page number.

**Galley:** A proof taken from composed type before page composition to allow for the detection and correction of errors.

**Gutter:** The inside margin of a publication and the space between two columns of type.

**Image area:** The area of a page inside the margins in which text and graphics are placed (objects may extend outside this area.)

**Legislative Editing XML Application (LEXA):** The customized version of XMetal currently used by the Senate for authoring bills.

**Locator code:** GPO tagging scheme that enables the specification of typesetting instructions into a data file.

**PostScript:** A page description language developed by Adobe Systems.

**Preliminary composition:** Preparatory representation of content format or structure.

**Publication settings and processing instructions:** Publication specific styling rules and instructions for page layout and typesetting.

**Unicode:** Unicode characters can be encoded using any of several schemes termed Unicode Transformation Formats (UTF). Unicode is an industry standard designed to allow text and symbols from all of the writing systems of the world to be consistently represented and manipulated by computers.

**UTF-8:** UTF-8 is a variable-length character encoding to represent Unicode characters.

**Workbench:** A customizable graphic user interface that allows users to interact with a

**FINAL**

software application through the selection of tools, key commands, help options, and other functionality.



FINAL

## **APPENDIX 2: ACRONYMS**

### **ASCII**

American Standard Code for Information Interchange

### **DTD**

Document Type Definition

### **EPS**

Encapsulated Postscript

### **GUI**

Graphical User Interface

### **HTML**

Hypertext Markup Language

### **HTTP**

Hypertext Transfer Protocol

### **JPEG**

Joint Photographic Experts Group

### **LEXA**

Legislative Editing XML Application

### **PDF**

Portable Document Format

### **SGML**

Standardized General Markup Language

### **TIFF**

Tagged Image File Format

### **WAIS**

Wide Area Information Server

### **W3C**

World Wide Web Consortium

### **XHTML**

Extensible Hypertext Markup Language

**FINAL**

**XML**

Extensible Markup Language

**XPATH**

XML Path Language

**XSL**

Extensible Stylesheet Language

**XSLT**

Extensible Stylesheet Language Transformations