

GP 1.2:D 56/SUM.

DALLAS PUBLIC LIBRARY
The Future Digital System - A W
GP 1.2:D 56/SUM.



X60101458

**The Future Digital
System- A world class
system for managing
official Government
content, which will
verify and track
versions, assure
authenticity, preserve
content, and provide
permanent access.**

GPO
Roadmap to
the Future

GOVERNMENT PUBLICATION
DEPOSITORY 594

DEC 13 2004

DALLAS PUBLIC LIBRARY



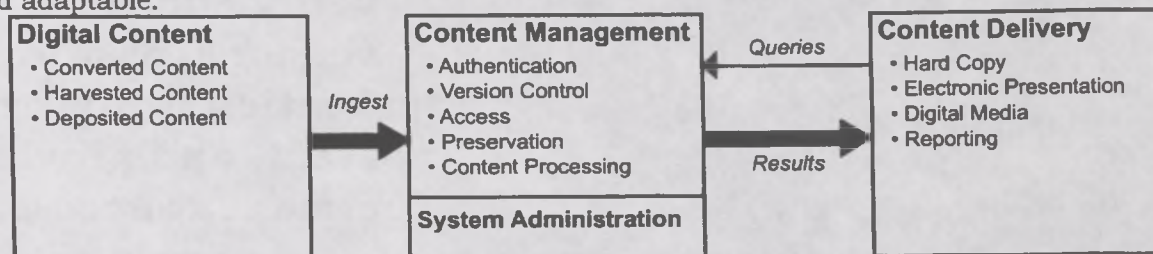
GPO

SETTING THE STAGE

A Concept of Operations (ConOps) document was created as a guideline for developing the Future Digital System. A ConOps document is the IEEE standard for describing the high level user requirements of a software-intensive system. This document concentrates on WHAT the system capabilities and requirements are, not HOW the system will be implemented. Based on the Open Archival Information System (OAIS) model, the Future Digital System will be a rules based, policy driven system capable of being configured to meet the needs of users. The ConOps is a reference model driven approach containing functional as well as user models. The ConOps document is currently in the review stage and will be available to the public at a future date.

Background

The Future Digital System will be a world class information life cycle management system that will transform GPO into a leadership content management agency. The Future Digital System will be system and policy neutral such that it can support changes resulting from technology, policy, or customer needs. The system will be modular and adaptable.



CRITICAL FUNCTIONS

- **Version Control:** Relating to a specific issuance, edition, or revision of a document or publication.

Version control within the Future Digital system will be the process of establishing the version of the content. Once the version has been established, the Future Digital System will evaluate and track it throughout its entire life cycle based on GPO established best practices.

The chain of responsibility will be reflected in the version information (ie., who created it, when, etc.)

- **Authentication:** Validation of a user, a computer, or some digital object to ensure that it is what it claims to be. In the specific context of the Future Digital System, the assurance that an object is as the author or issuer intended it.

The Authentication function will verify that digital content within the Future Digital System is authentic or official, and certify this to users accessing the content *as needed*. When a digital publication is received by GPO, the Future Digital System must have the capability to verify the authenticity of the content and determine its status as official content.

The Future Digital System will have the capability to provide users who access content with a token (or digital certificate) containing a visual representation that the content is certified to be official or authentic. The Future Digital System will also provide a means by which sections or small pieces of a document could be digitally certified. This concept is otherwise known as granularity.

- **Preservation:** Maintaining information products over time for use in their original form or in some other verifiable, usable form.

GPO's users expect us to provide permanent public access to official U.S. Government publications. GPO's responsibility for this is mandated by law. Preservation storage for current and permanent access are key elements of the Future Digital System:

The active storage component, or access archive, includes files used to deliver content to users.

Preservation archival storage, the Collection of Last Resort (CLR), is a permanent collection of the final published, official version(s) of U.S. Government publications in all formats intended to provide comprehensive, timely, permanent public access.

The CLR will consist of multiple collections of tangible and digital publications with identical content located at geographically-separated sites. This redundancy ensures continuity of access and guaranteed permanence in the event of a disaster or significant discontinuation of service at a single site. Creation of the CLR will follow best practices in the preservation community. The exact number of sites is yet to be determined. There is no direct public access or use of content in the Collection of Last Resort. Only when all other avenues of access are exhausted does the official content in the CLR come into use. Then access copies, called derivatives, are produced to support public access.

- **Access:** The primary interface between End Users and the the Future Digital System.

The 5 major components of Access include Finding, Ordering, Retrieval, User Support, and Data Mining.

Finding is comprised of three tools that interact seamlessly to create a total user experience:

Search Tools perform queries on content and metadata.

Reference Tools are lists and resources that point to content.

National Bibliography (or Cataloging) Tools add descriptive metadata to content in the form of library standard bibliographic records.

Ordering provides the capability for users to place orders and submit payment electronically.

Retrieval Tools provide retrieval instructions so that content may be retrieved from storage and ultimately delivered to End Users.

User Support Tools assist in delivering services to End Users. User Support tools may include alert services, CRM tools, knowledge bases, and stored user preferences.

Data Mining Tools find, aggregate, and associate business process information. This type of information may include document retrievals and agency billing information.

CONTENT

Digital Standards

Standards are the means for our participation in the communities we serve, making information available and usable. The Future Digital System will be driven by GPO developed best practices that support Government publishing. This includes the creation of:

- **Metadata:** Data that describes the content, quality, condition, or other characteristics of other data.

Metadata describes how, when, and by whom information was collected, where it resides, and how it is formatted. Metadata helps locate, interpret, and manage content throughout its lifecycle in the Future Digital System. Virtually every function within the Future Digital System will require, and often create, information describing the context, quality, condition, use, or nature of content, users, or processes. The system will require tools to create, interpret, and share this metadata between and among a wide variety of existing frameworks and standards already deployed in the various communities of use and practice whom we serve.

- **Preservation Digital Master:** A faithful, high quality version of content which is created in a controlled environment, bound together with information which supports long-term preservation.

We will, with awareness of and leadership in the developing best practice for digital preservation, define specification for a uniform package of digital information to serve as a Preservation Digital Master. It will capture, at the highest possible quality, digital content intended for preservation in the Future Digital System, together with rich metadata, which will support preservation processes over the long term. The goal of that preservation is a contextually meaningful, fully functional, and complete rendering of the content which can be certified as official and faithful to the original issue, regardless of obsolescences and other shifts in technology.

- **Style Tools:** Tools that help customers create and provide content to the Future Digital System.

Style tools will allow GPO to move further upstream in the publishing process - towards content creators - which may allow us to fully capture content. Style Tools are designed to validate to the proposed new GPO Style Manual.

The Future Digital System will contain 3 possible style tools:

A Validator assures that content is acceptable for the Future Digital System (based on GPO best practices)

Conversion tools that convert content in structured or proprietary formats (e.g., Quark) to a package that is accepted by the Future Digital System

Auto-Composition tools that provide a comprehensive range of services to agencies (Auto Composition, Collaboration, etc.)

Digital Content

The focus of the Future Digital System will be on born digital content. The Future Digital System will also support converted legacy content and harvested material.

- **Deposited Content:** Electronic content that is pushed to the Future Digital System by originating agencies for preservation and access.

Based on the ease of system usage and providing Style Tools, it will streamline the process for customers using the Future Digital System. Deposited Content will be consistent with the digital standards for the system. The system will not be limited to the formats that exist today.

• **Harvested Content:** Electronic documents that are first published directly to agency Web sites and then pulled into the the Future Digital System in a way that is consistent with digital standards.

It will be necessary for the Future Digital System to use a suite of tools to discover, harvest, and assess content.

Discovery Tools will be used to discover electronic documents that are in scope.

Harvesting Tools will systematically capture content and available metadata.

Assessment Tools will be used to determine if a document is in scope and if a version of that document already exists in the system.

If the document is in scope and not already in the Future Digital System, it will be ingested into the system.

• **Converted Content:** Electronic files created from tangible documents which can then be preserved and derived into new digital products.

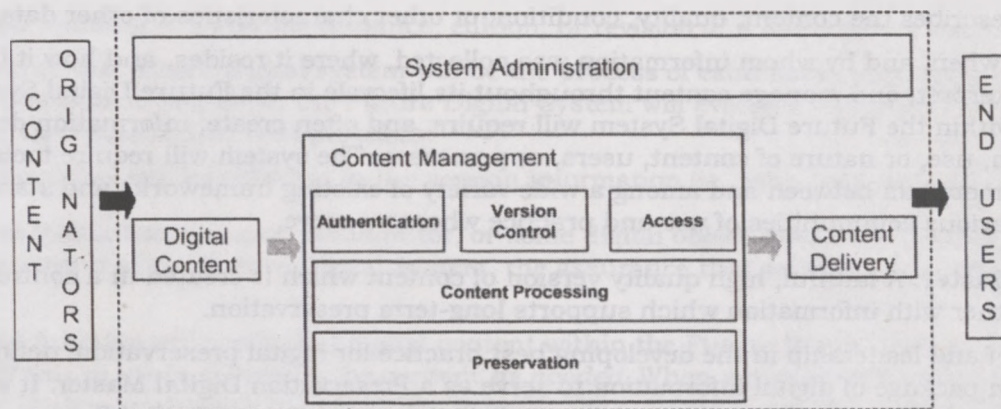
GPO will work with various user communities (including the library community) on digitizing the collection of legacy documents held in Federal depository libraries. This collection will be made available in the public domain for permanent public access through GPO's dissemination programs.

Conversion must be performed at a level that is adequate to support preservation as well as public access.

As a result, converted files must be compatible with standards and best practices for conversion and digital preservation masters.

CONTENT MANAGEMENT

The Future Digital System's working layer that controls and manages content.



CONTENT DELIVERY

The Future Digital System will deliver content to meet user needs and in conformance with GPO developed best practices. This includes, but is not limited to the following delivery methods:

- Hard Copy is tangible printed content. Print on Demand is hard copy produced in a short production cycle time and typically in small quantities.
- Electronic Presentation is defined as a delivery mechanism of the dynamic and temporary representation of content in digital format.
- Digital Media is a content delivery mechanism consisting of data storage devices. Digital media includes:
 - Data storage devices
 - Wireless handheld devices
 - Future media (e.g., flexible electronic displays, etc.)
 - Storage at user sites (e.g. subscription services, etc.)

As electronic presentation becomes the primary method of disseminating Government publications, the goal will shift away from making the digital format match the printed product and will instead focus on presenting the electronic content in the most useable format.

Regardless of format, the primary goal of delivery is to meet the requirements of users. The Future Digital System will provide capabilities based on those requirements. Tools will allow users to customize, compile, and proof content before delivery.