Agenda

U.S. GOVERNMENT PRINTING OFFICE

October 6, 2005

Check In

8:30 AM	**Vendor Check-In**
9:25 AM	**Check-in Closes**

Segment 1

9:30 AM	**Welcome and Keynote**
9:50 AM	**FDsys Overview** • System Clusters • Release Schedule
10:05 AM	**FDsys Acquisition** • Strategy Outline • Schedules
10:20 AM	**Escorts to Breakout Session 1**

Segment 2

10:35 AM	**Interactive Breakout Session 1** Breakout 1 – Master Integrator Breakout 2 – Content Preservation Breakout 3 – Content Submission Breakout 4 – Infrastructure Breakout 5 – Content Management Breakout 6 – Content Delivery Breakout 7 – Content Access
11:20 AM	**Escorts to Breakout Session 2**
11:35 AM	**Interactive Breakout Session 2** Breakout 1 – Master Integrator Breakout 2 – Content Preservation Breakout 3 – Content Submission Breakout 4 – Infrastructure Breakout 5 – Content Management Breakout 6 – Content Delivery Breakout 7 – Content Access
12:20 PM	**Escorts to Main Staging Area**

Close

12:30 PM	**Escorts out of GPO**

After Industry Day

What Should You Do?

1) Write Capability Statements and Assessment Documents

Capability Statements: Keeping GPO’s goals in mind please outline how your company would perform the services of a Master Integrator (MI) or how your company’s products, services and/or technology map to one or more of the six solution clusters (e.g., the Industry Day breakout sessions).

Vendors interested in the role of MI are hereby requested to submit a capability statement (10 pages or less and no marketing material) that addresses their ability to perform the services of Master Integrator. Vendors interested in specific system clusters may submit capability statements (10 pages or less and no marketing material) that address one or more of the six solution clusters. Vendors may submit statements for MI and separate statements for any or all of the solution clusters.

Assessment Documentation: In addition, vendors are hereby encouraged to submit assessment documents (e.g., white papers or other documentation - no more than 50 pages) that address any gaps or concerns with the proposed acquisition, implementation plan and/or the FDsys documentation. Feedback should include specific comments and suggestions for refining the material to reflect clear guidance to industry and industry best practices.

NOTE: Please also list company points of contact and GSA Schedule number (if applicable).

2) Questions for the FDsys PMO Team

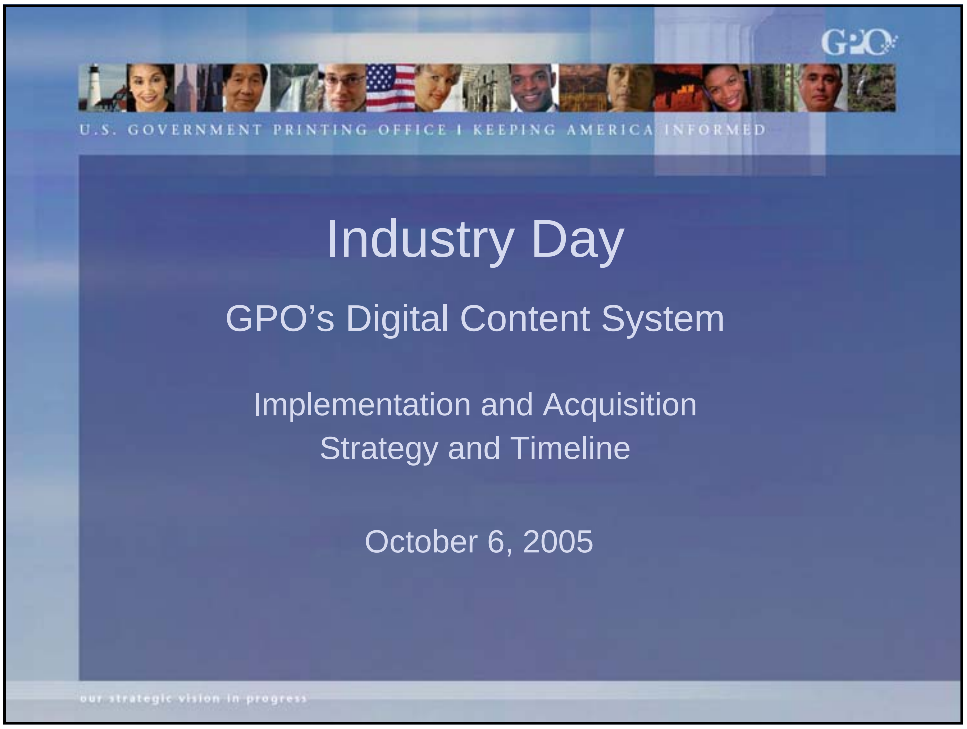
Questions for the PMO team or questions regarding capability statements, white papers and other documentation must be submitted via email (with the subject line **”Question about FDsys Industry Day”** to Herb Jackson at hjackson@gpo.gov. Questions must be submitted by noon on 13 October 2005. Questions submitted after this date and time will not be answered.

3) Submitting Material

Capability statements and/or other white papers and documentation related to Industry Day must be submitted by noon on 20 October 2005, to hjackson@gpo.gov. Documents must be submitted as **Adobe PDF or Microsoft Word files (2000 or higher)**. Faxed copies are not acceptable.

Notice

The industry day event and related RFI are requests for information only and do not obligate the GPO in any way. The event and subsequent dialog are not a request for proposal and the GPO will not pay for any information submitted or for any expenses associated with providing information. Any information submitted by respondents to the event and the RFI is strictly voluntary. Material submitted will be deemed proprietary to the extent permitted by applicable laws and regulations if so marked by the respondent.



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

GPO

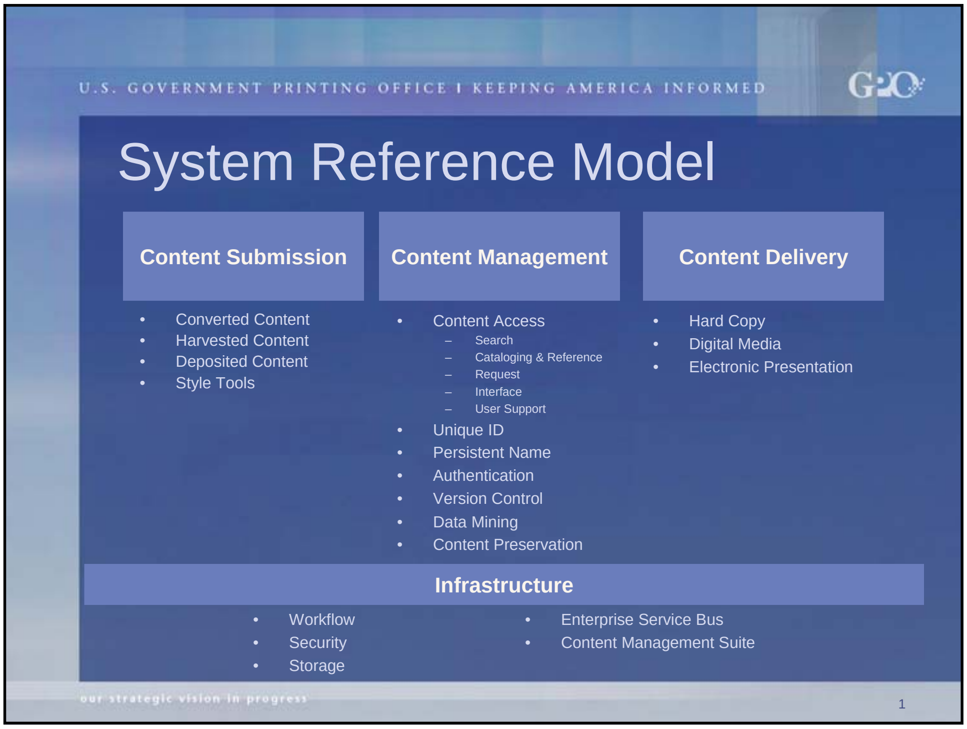
Industry Day

GPO's Digital Content System

Implementation and Acquisition
Strategy and Timeline

October 6, 2005

our strategic vision in progress



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

GPO

System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

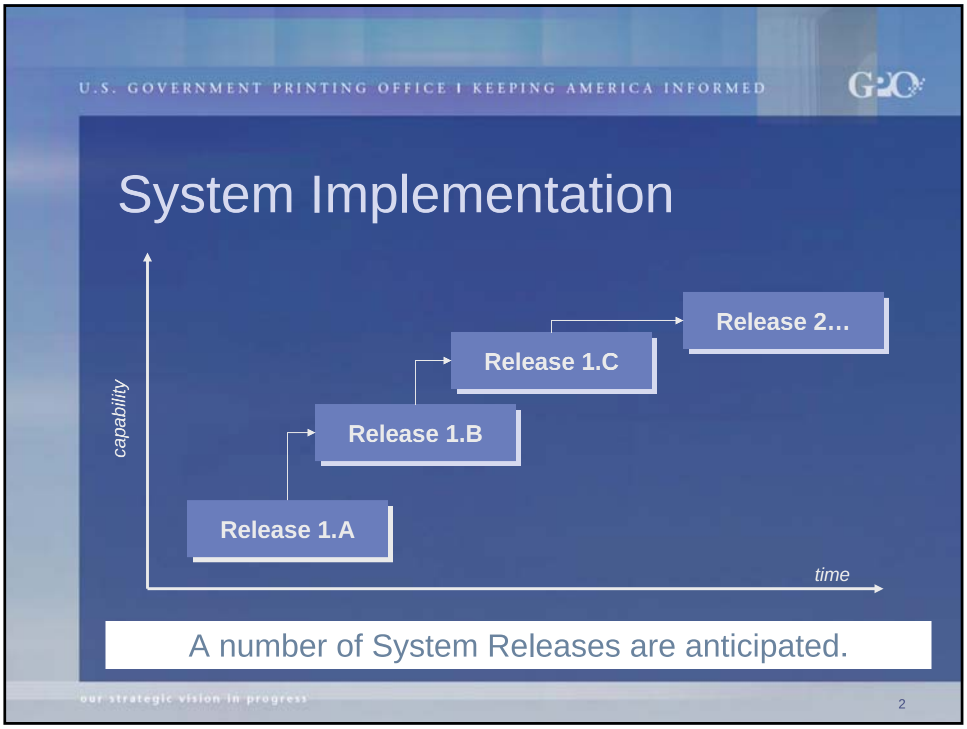
Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite

our strategic vision in progress



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

GPO

System Implementation

capability

Release 1.A

Release 1.B

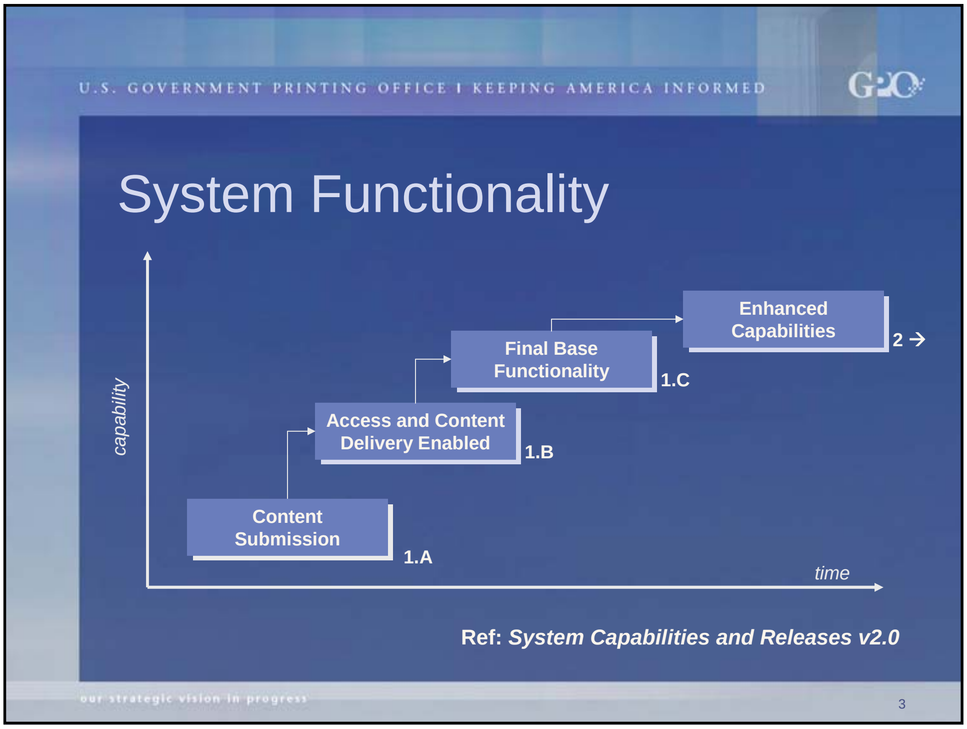
Release 1.C

Release 2…

time

A number of System Releases are anticipated.

our strategic vision in progress



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

GPO

System Functionality

capability

Content Submission 1.A

Access and Content Delivery Enabled 1.B

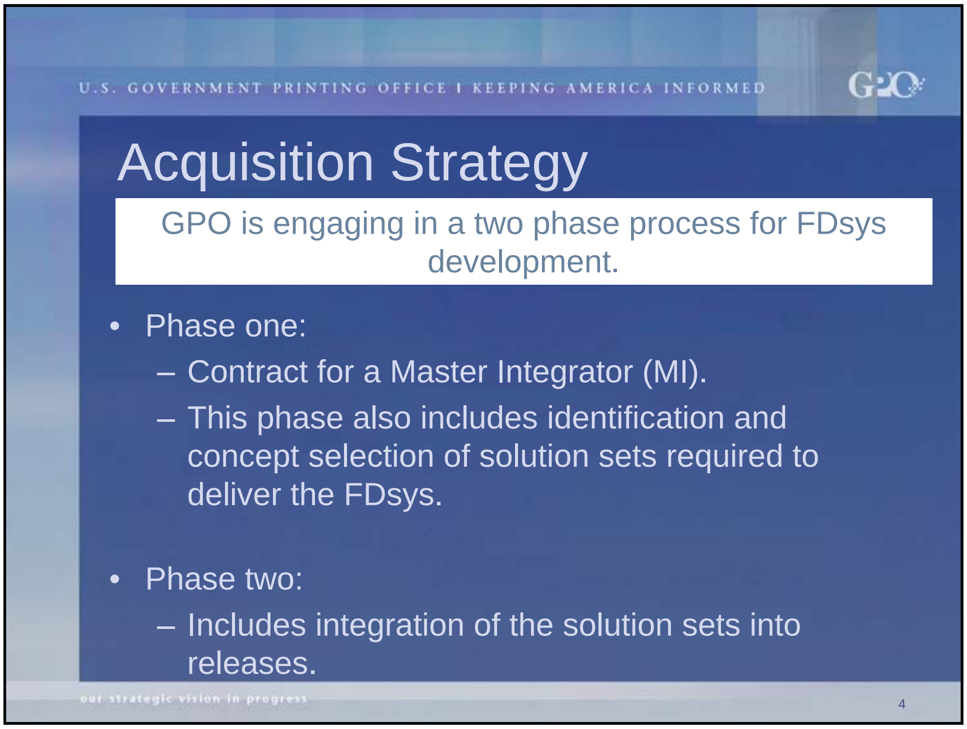
Final Base Functionality 1.C

Enhanced Capabilities 2 →

time

Ref: *System Capabilities and Releases v2.0*

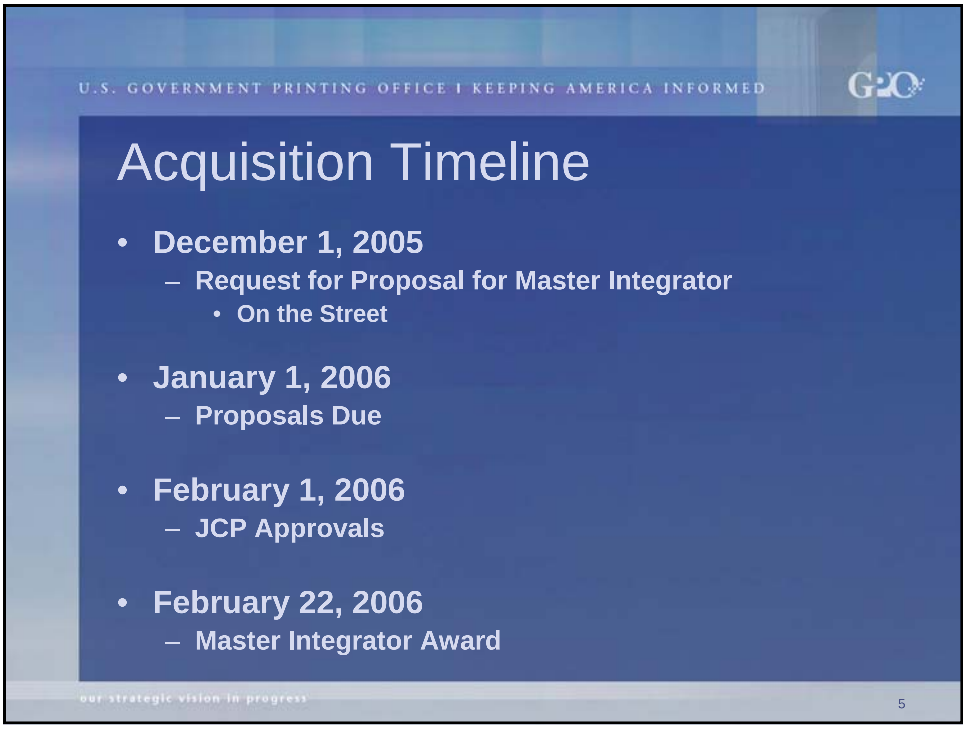
our strategic vision in progress



Acquisition Strategy

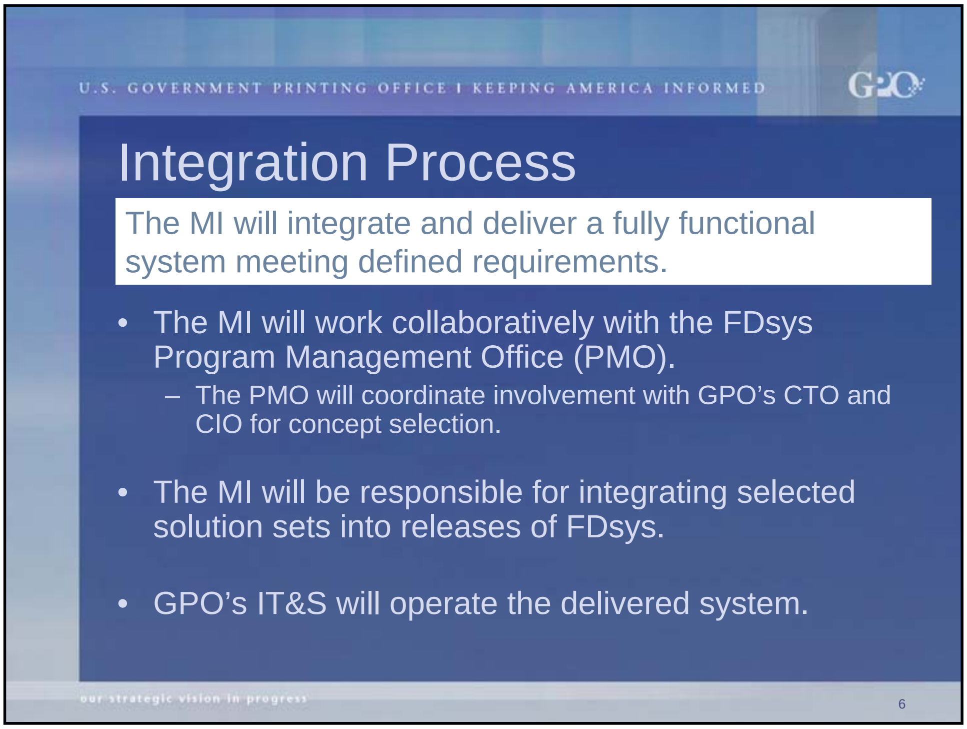
GPO is engaging in a two phase process for FDsys development.

- Phase one:
  - Contract for a Master Integrator (MI).
  - This phase also includes identification and concept selection of solution sets required to deliver the FDsys.
- Phase two:
  - Includes integration of the solution sets into releases.



Acquisition Timeline

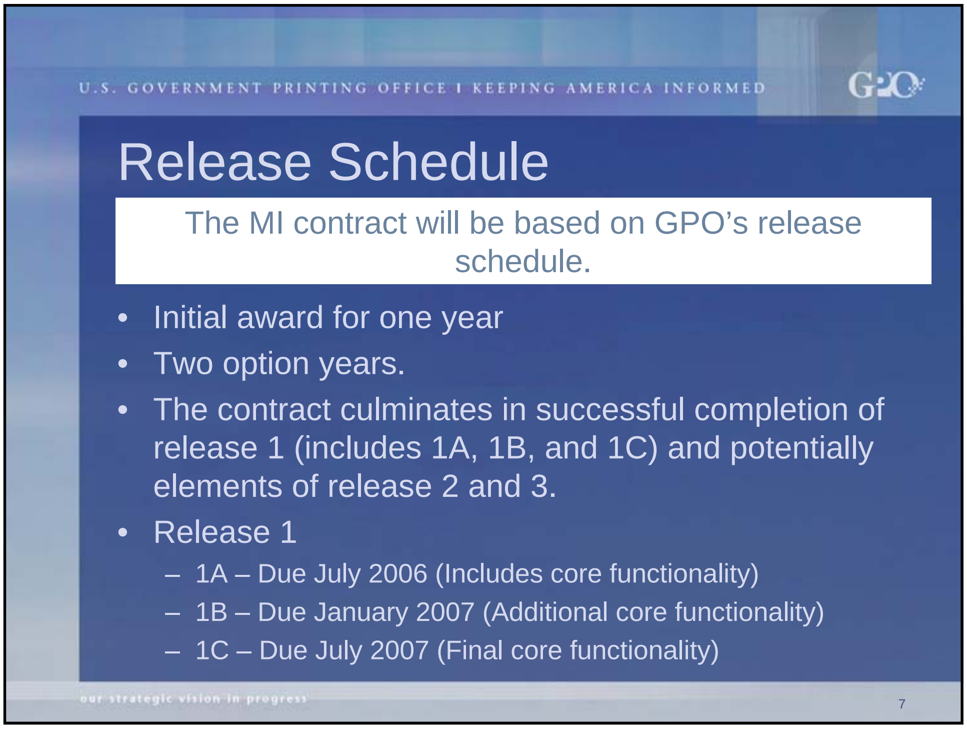
- **December 1, 2005**
  - **Request for Proposal for Master Integrator**
    - **On the Street**
- **January 1, 2006**
  - **Proposals Due**
- **February 1, 2006**
  - **JCP Approvals**
- **February 22, 2006**
  - **Master Integrator Award**



Integration Process

The MI will integrate and deliver a fully functional system meeting defined requirements.

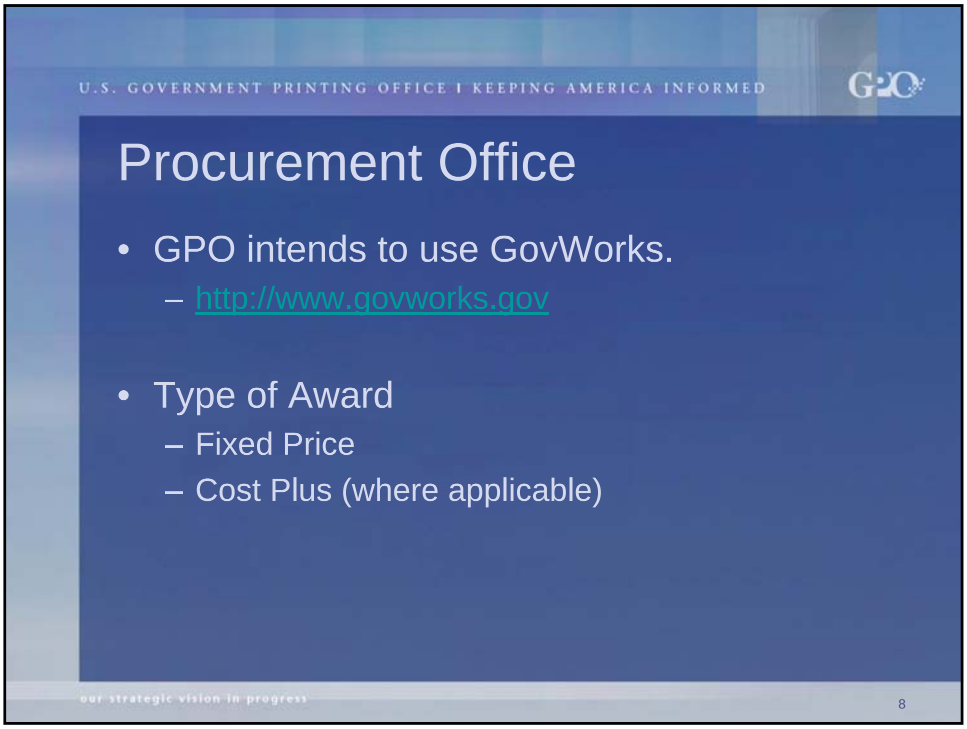
- The MI will work collaboratively with the FDsys Program Management Office (PMO).
  - The PMO will coordinate involvement with GPO’s CTO and CIO for concept selection.
- The MI will be responsible for integrating selected solution sets into releases of FDsys.
- GPO’s IT&S will operate the delivered system.



Release Schedule

The MI contract will be based on GPO’s release schedule.

- Initial award for one year
- Two option years.
- The contract culminates in successful completion of release 1 (includes 1A, 1B, and 1C) and potentially elements of release 2 and 3.
- Release 1
  - 1A – Due July 2006 (Includes core functionality)
  - 1B – Due January 2007 (Additional core functionality)
  - 1C – Due July 2007 (Final core functionality)



Procurement Office

- GPO intends to use GovWorks.
  - http://www.govworks.gov
- Type of Award
  - Fixed Price
  - Cost Plus (where applicable)

After MI - Next Steps

- Independent Validation and Verification (IV&V)

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 1

Master Integrator

In GPO's FDsys acquisition strategy, the Master Integrator (MI) is a prime contractor to implement the FDsys and integrate system components selected by GPO.

MI responsibilities include:

- Development of all integration needs
- Acquisition of the key system components, technologies and applications.
  - As collaboratively selected by GPO and the MI.
- Assistance
  - Testing
  - Training
- Conversion of or integration of legacy systems and their functionality as described in the to be developed Enterprise Architecture and the provided TRM
  - Loading of content and other initial populating of FDsys' content.
- Operations (24 x 7)
  - Through beta
  - Including system testing
- Provide all necessary plans and participate in GPO reviews

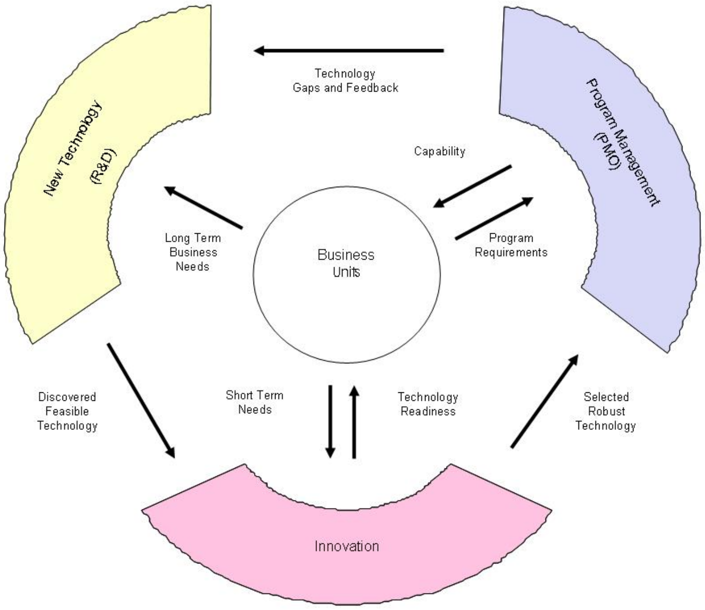
Master Integrator:

Master Integrator (MI) is GPO’s term for a systems integrator. For the FDsys, the MI will be an individual or company that unites various components, technology and applications of the FDsys functional clusters into a world-class information lifecycle management system. The MI will work collaboratively within GPO’s Technology Management structure to select the components of technology that meet or exceed defined requirements.

- The MI will work collaboratively with the FDsys Program Management Office (PMO), Chief Technical Officer and Chief Information Officer, and their staff, to perform concept selection and reliability and robustness testing of the solution sets for clusters identified by GPO.

- Once the selection of technology is accomplished, the MI will be responsible for integrating those solution sets into the subsequent working releases of FDsys. The ultimate objective for the MI will be to integrate and deliver a fully functional system to meet the requirements as defined in the FDsys documentation.

- The MI will also assist GPO with system testing and training, and will provide initial operating capability through beta.



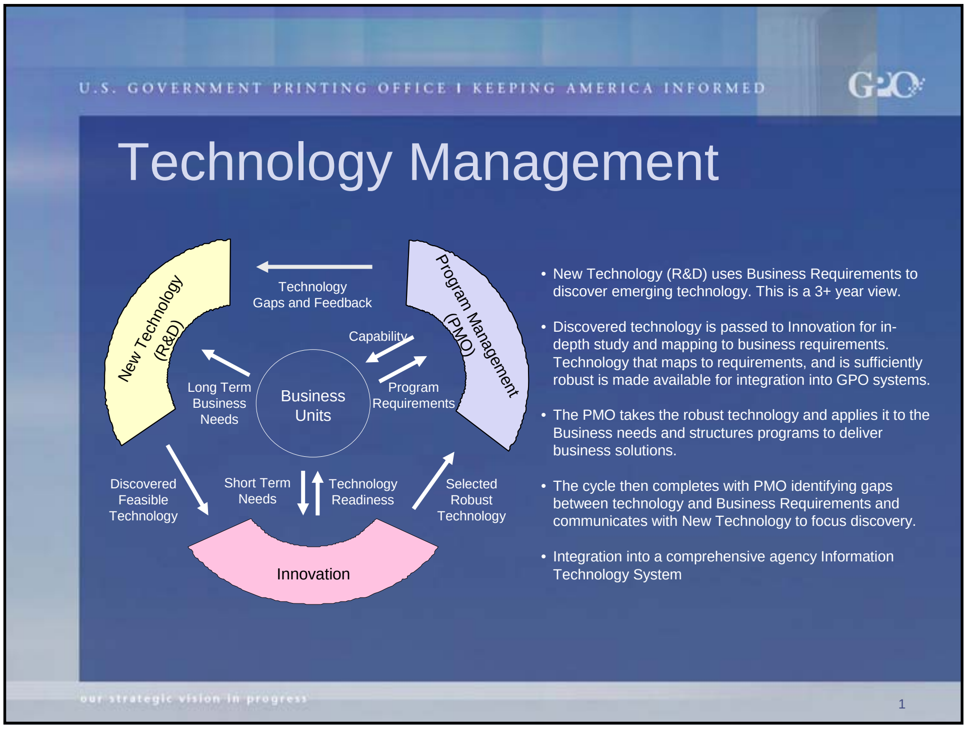
GPO’s FDsys will be a complex, software intensive system comprised of over 20 functional elements. Each of these 20 functional elements is individually complex from a systems perspective, and will consist of hardware, software and data components. After 2 years of exhaustive market research, planning, concept design, and requirements writing, GPO is prepared to move forward developing the FDsys.

Industry Day

GPO's Master Integrator Needs

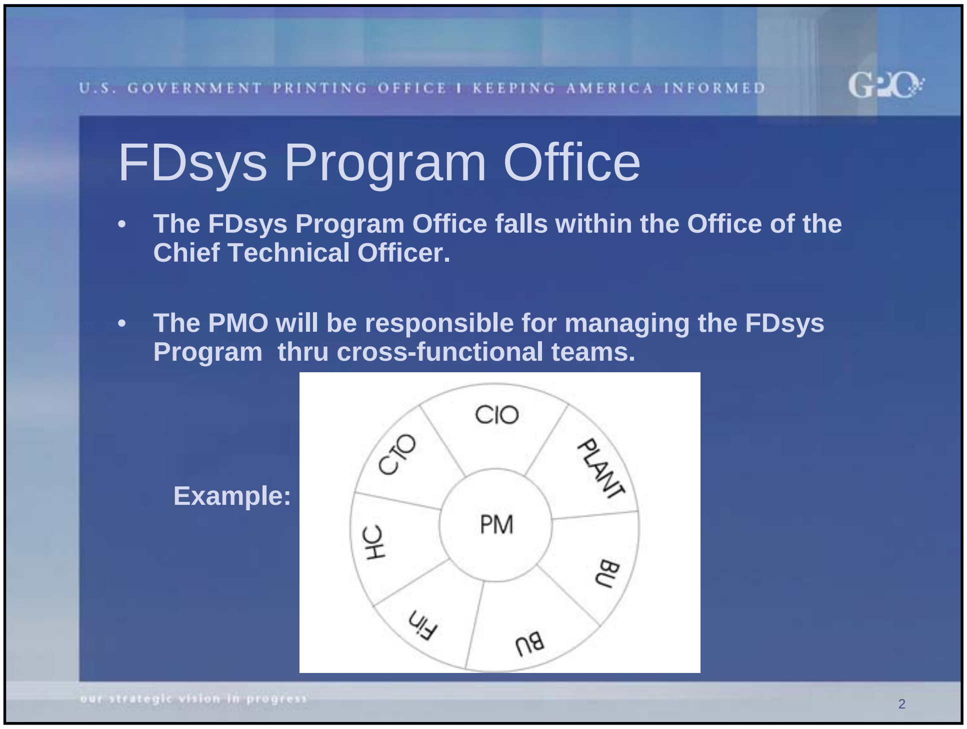
Breakout Session 1

October 6, 2005



Technology Management

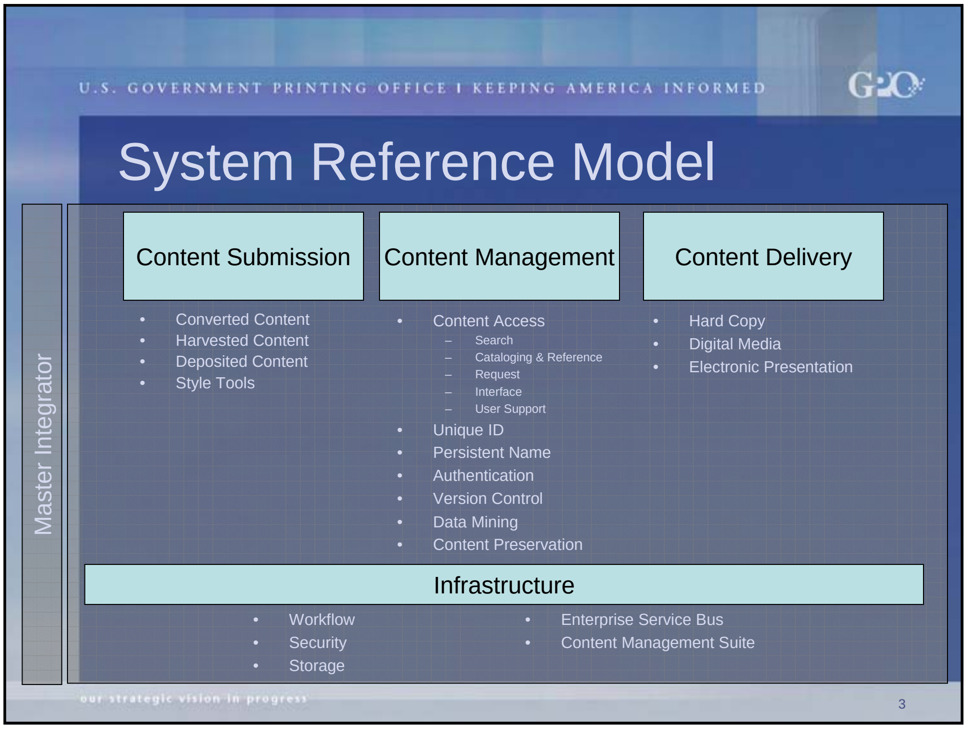
- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System



FDsys Program Office

- The FDsys Program Office falls within the Office of the Chief Technical Officer.
- The PMO will be responsible for managing the FDsys Program thru cross-functional teams.

Example:



System Reference Model

Master Integrator

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

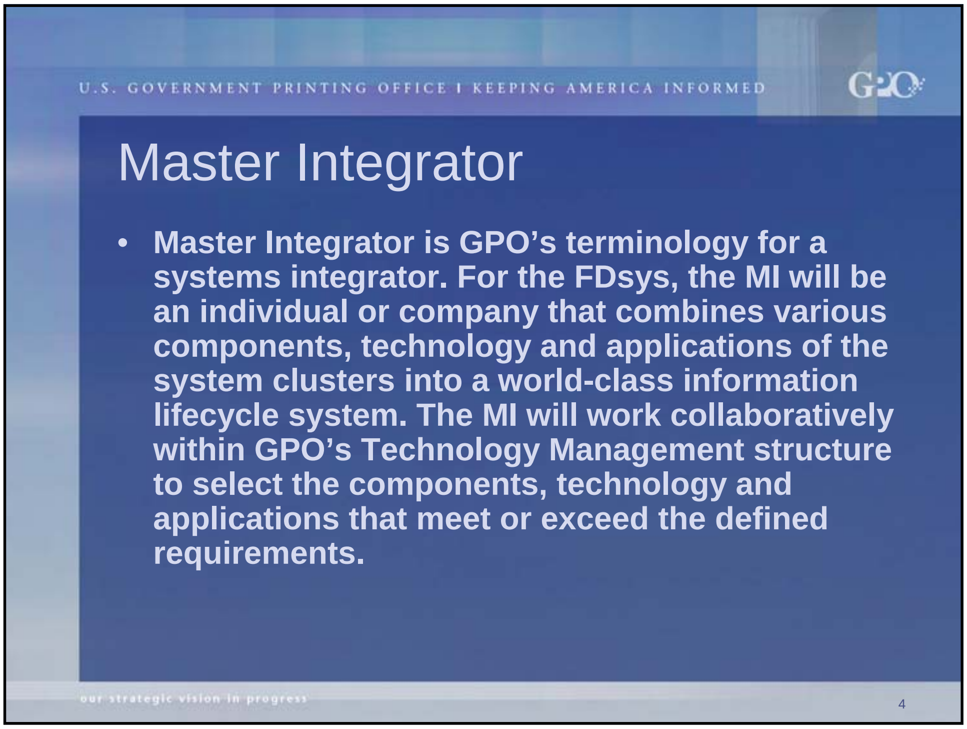
- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

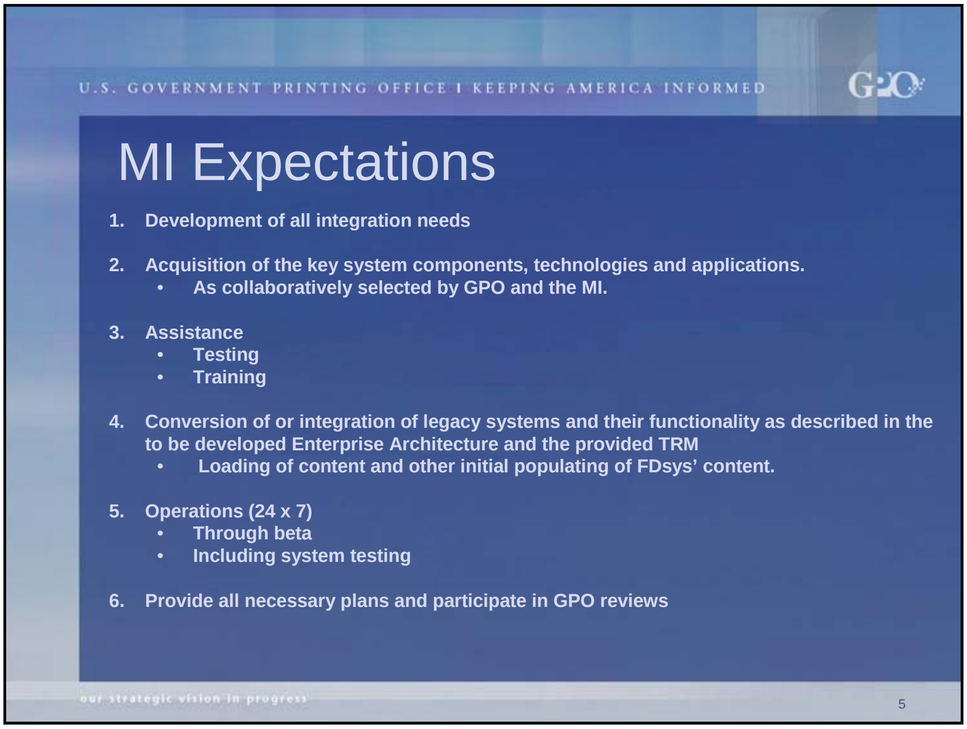
Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite



Master Integrator

- Master Integrator is GPO's terminology for a systems integrator. For the FDsys, the MI will be an individual or company that combines various components, technology and applications of the system clusters into a world-class information lifecycle system. The MI will work collaboratively within GPO's Technology Management structure to select the components, technology and applications that meet or exceed the defined requirements.



MI Expectations

1. Development of all integration needs

2. Acquisition of the key system components, technologies and applications.
   - As collaboratively selected by GPO and the MI.

3. Assistance
   - Testing
   - Training

4. Conversion of or integration of legacy systems and their functionality as described in the to be developed Enterprise Architecture and the provided TRM
   - Loading of content and other initial populating of FDsys' content.

5. Operations (24 x 7)
   - Through beta
   - Including system testing

6. Provide all necessary plans and participate in GPO reviews

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 2

Content Preservation

FDsys Content Preservation enables comprehensive, timely, permanent public access to the final published, official versions of U.S. Government publications in digital formats, by the retention of faithful, fully functional master files of content, and the performance of processes which assure ongoing usability of those files.

Preservation processes include activities associated with maintaining digital publications for use, either in their original form or in some verifiable, usable form. Preservation may include creation of a surrogate for the original by a transformation process, wherein the intellectual content and other essential attributes of the original are retained. Digital preservation processes may include refreshment, migration, emulation, and other emerging techniques.

Breakout Session White Paper

U.S. GOVERNMENT PRINTING OFFICE

Content Preservation

Background

Preservation enables comprehensive, timely, permanent public access to the final published, official version(s) of U.S. Government publications in digital formats, by the retention of faithful, fully functional master files of content, and the performance of processes which assure ongoing usability of those files. Preservation copies of digital publications, Archival Information Packages (AIPs), with associated technical metadata, will be maintained in Future Digital System Archival Storage.

Inputs and Outcomes

Preservation **inputs** are Archival Information Packages (AIPs), which are currently executable digital files of all types, with their behaviors controlled by applications.

In order of preference, the **outcomes** desired are:

- Faithfully duplicated files, rendered using the original application.
- Files which faithfully reproduce content, behavior and appearance of the original, rendered using other software than the original application.
- Files which exactly convey the content but may alter behavior and/or appearance, rendered using other software than the original application.

How are these outcomes achieved?

- **Refreshment** (copying) of content to new media. Refreshment is defined as a preservation process for data extraction, cleaning and integration, and the triggering events of these activities.
- **Migration** of files from compatibility with one application or format to subsequent versions of that application or format. Migration includes modifying files from compatibility with one application or format to a version which can be indefinitely maintained. Migration is defined as preservation of digital content where the underlying information is retained but older formats and internal structures are replaced by newer.
- **Emulation** of essential behaviors and attributes of digital objects by characterization based on metadata. Emulation is defined as replication of a computing system to process programs and data from an earlier system that is no longer is available.
- **Hybrids** of these approaches, or new approaches.

Preservation Process Triggers

Preservation processes are triggered by an assessment or a user action. Assessment criteria for initiating a process include:

- Scheduled event (generally results in Refreshing)
- Application failure (Total loss of functionality)
- System-detected loss of content, functionality, or metadata
- Managed request (from a Service Specialist)
- Request for new type of derivative for access
- Scheduled random sampling of content in AIP Storage

File Management and Disposition

In addition to the preservation processes outlined above, GPO also requires other supporting activities necessary to keep content accessible and usable, including:

- Management of preservation processes
- A secure repository environment including secure storage, file backup and redundancy
- File disposition options

U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

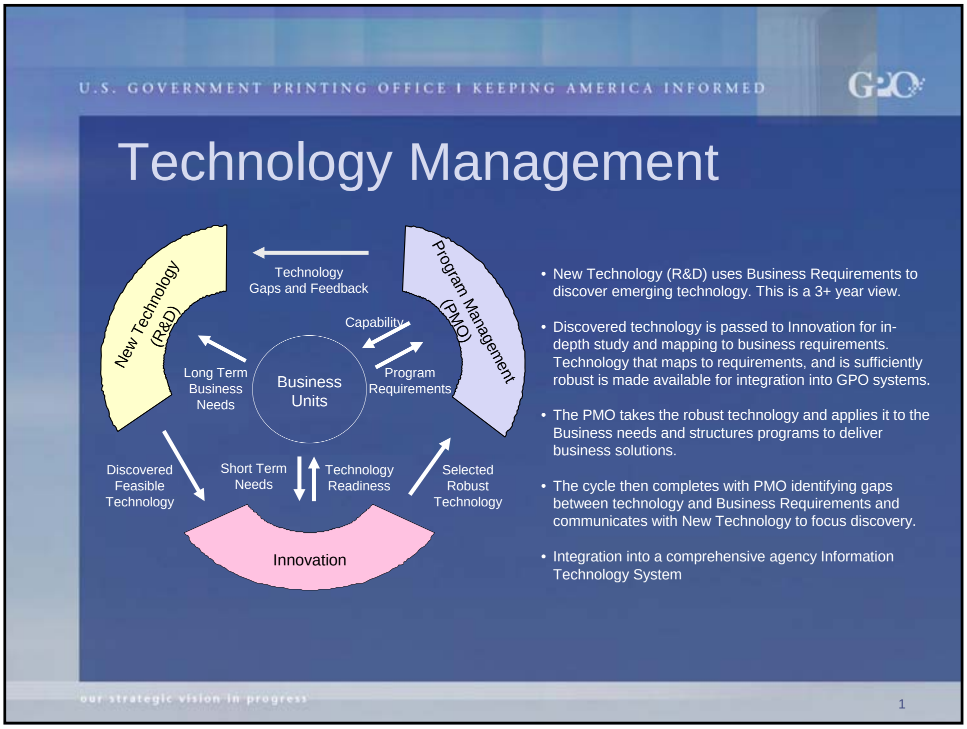
Industry Day

Content Preservation

Breakout Session 2

October 6, 2005

our strategic vision in progress

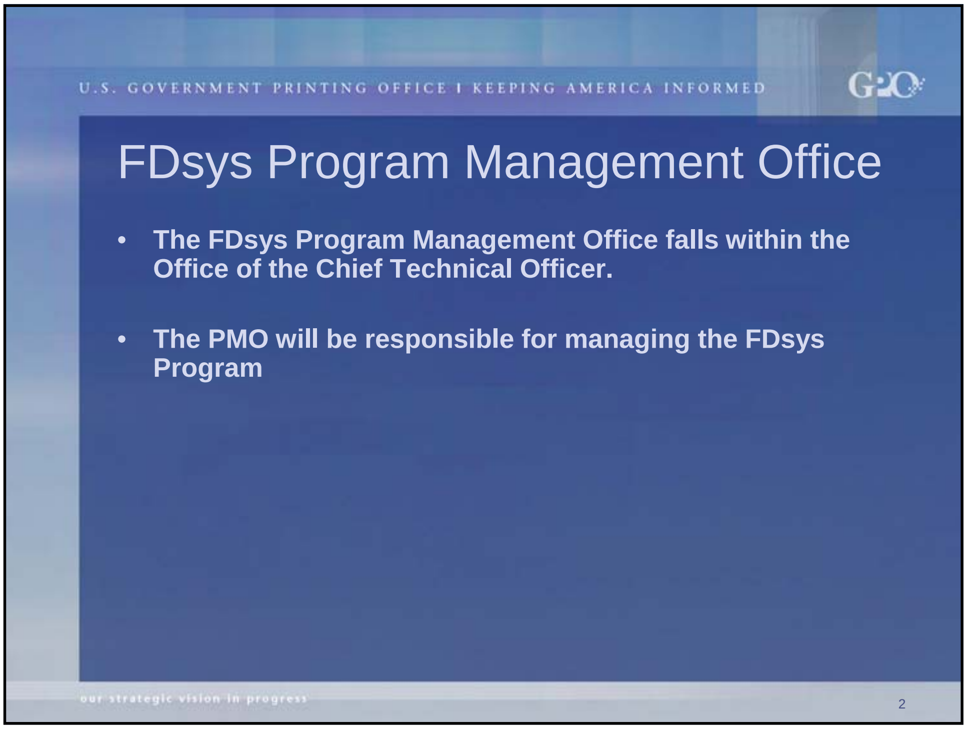


U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

Technology Management

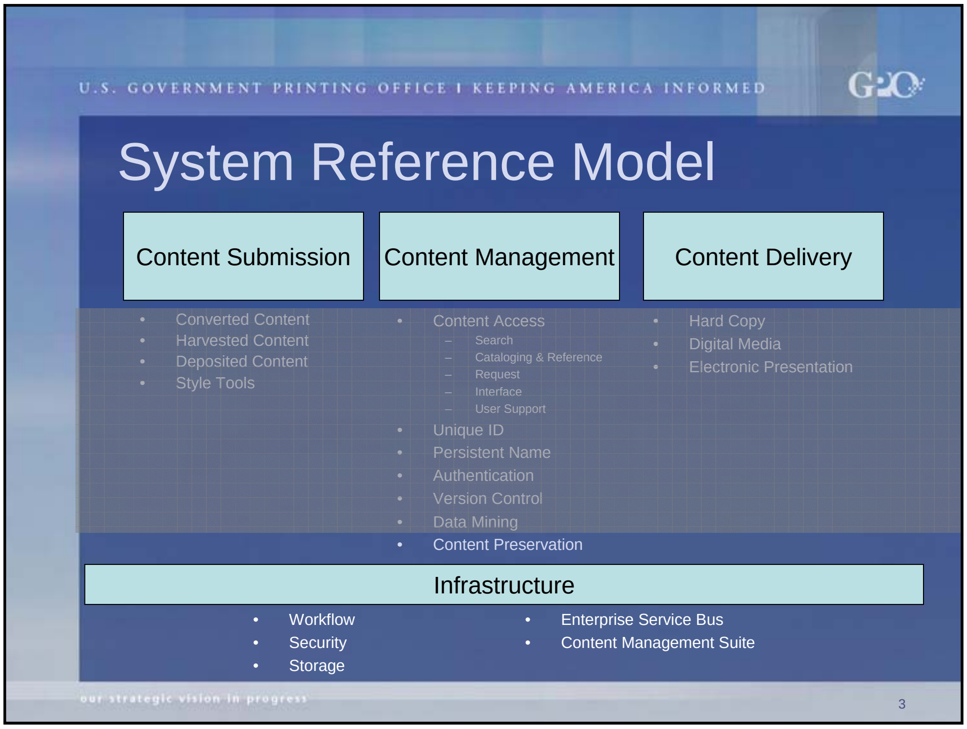
- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System

our strategic vision in progress



FDsys Program Management Office

- **The FDsys Program Management Office falls within the Office of the Chief Technical Officer.**
- **The PMO will be responsible for managing the FDsys Program**



System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

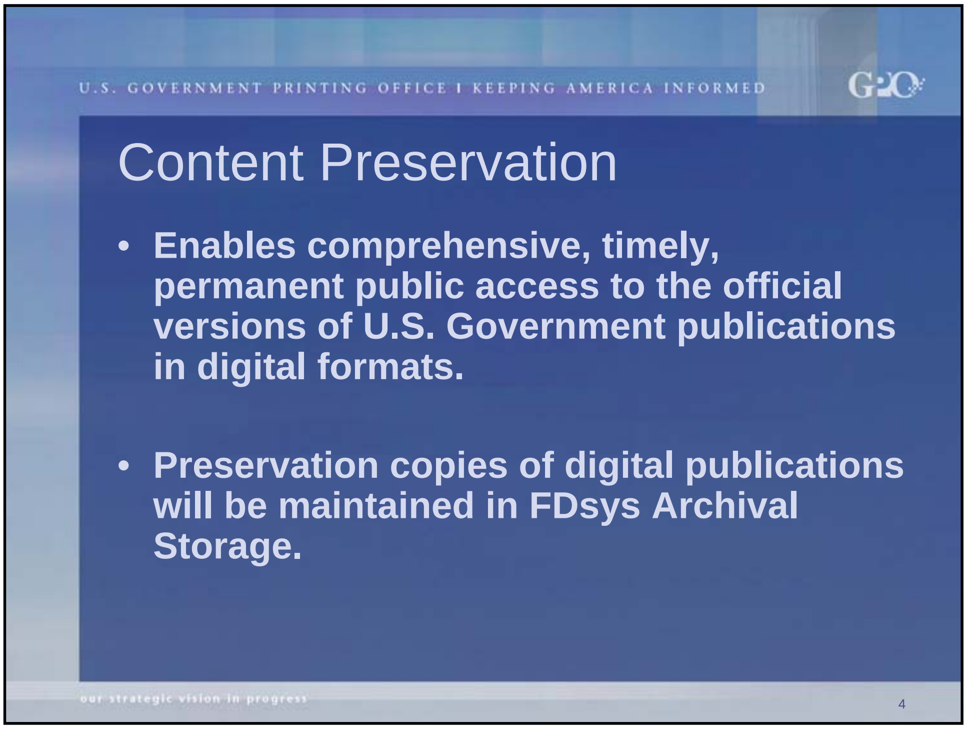
- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

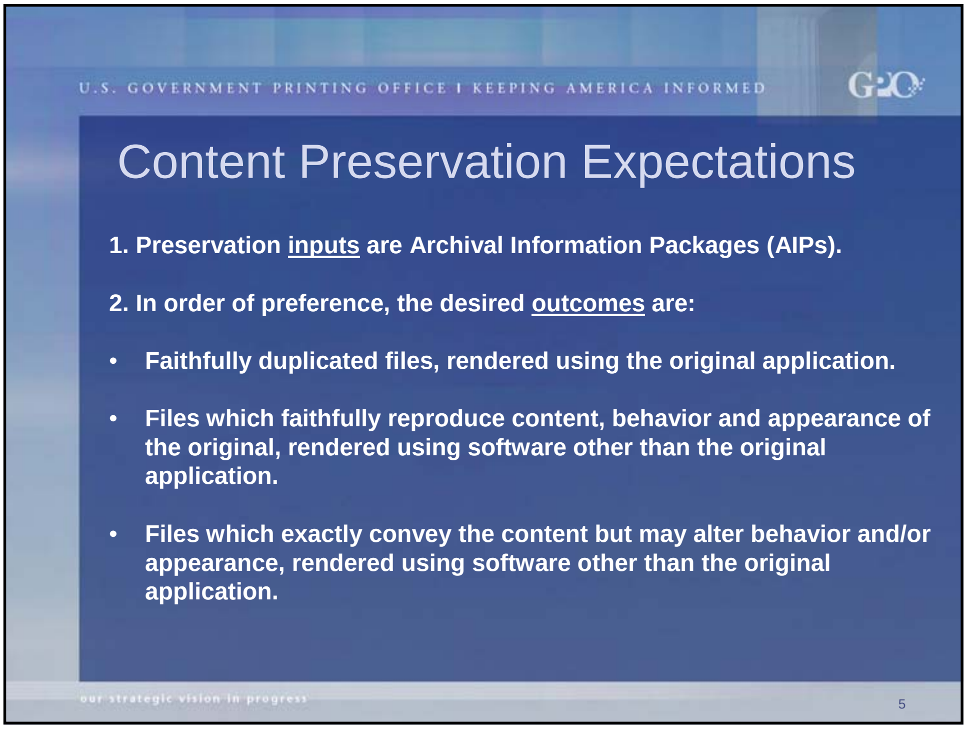
Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite



Content Preservation

- Enables comprehensive, timely, permanent public access to the official versions of U.S. Government publications in digital formats.
- Preservation copies of digital publications will be maintained in FDsys Archival Storage.

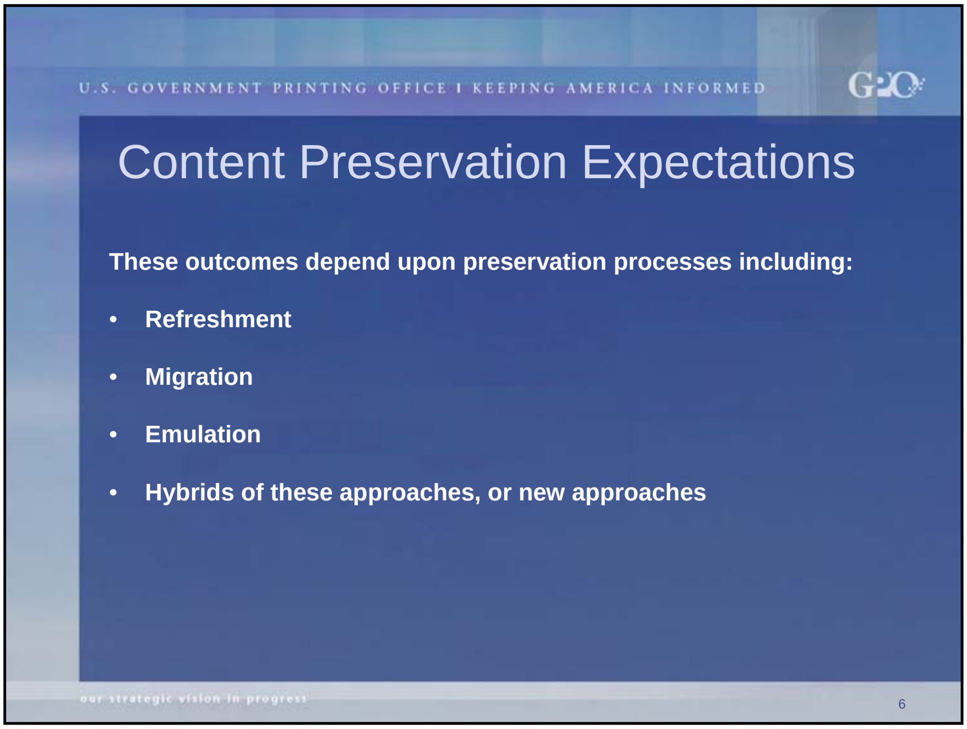


Content Preservation Expectations

1. Preservation inputs are Archival Information Packages (AIPs).

2. In order of preference, the desired outcomes are:

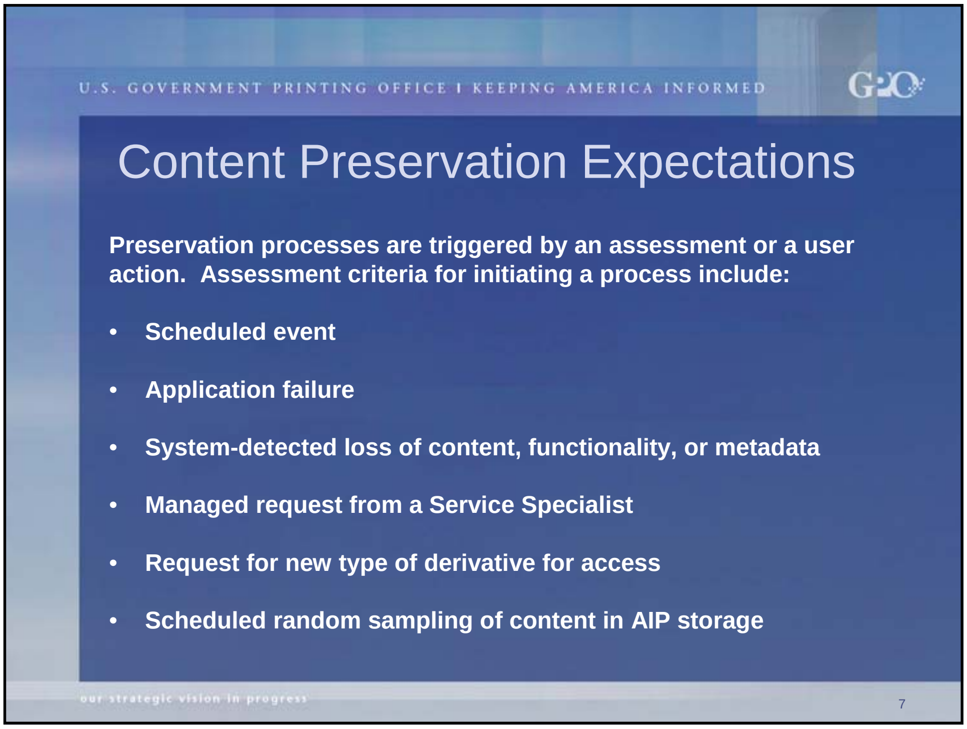
- Faithfully duplicated files, rendered using the original application.
- Files which faithfully reproduce content, behavior and appearance of the original, rendered using software other than the original application.
- Files which exactly convey the content but may alter behavior and/or appearance, rendered using software other than the original application.



Content Preservation Expectations

These outcomes depend upon preservation processes including:

- **Refreshment**
- **Migration**
- **Emulation**
- **Hybrids of these approaches, or new approaches**



Content Preservation Expectations

Preservation processes are triggered by an assessment or a user action. Assessment criteria for initiating a process include:

- **Scheduled event**
- **Application failure**
- **System-detected loss of content, functionality, or metadata**
- **Managed request from a Service Specialist**
- **Request for new type of derivative for access**
- **Scheduled random sampling of content in AIP storage**

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 3

Content Submission

The Content Submission cluster includes the processes to bring digital content into the FDsys and to enable Content Originators to order GPO services through FDsys.

The three types of content are:

- Deposited content: content intentionally submitted to GPO by Content Originators.
- Harvested content: content within the scope of dissemination programs that is gathered from Federal agency Web sites.
- Converted content: digital content created from a tangible product.

Style Tools will allow Content Originators to prepare content prior to ingest by GPO by providing tools for content capture, composition, collaboration, and approval.

Content Originator (CO) Ordering provides a system interface for Content Originators. Content Originators may submit content, order and re-order content, and specify Content Delivery and other service options through CO Ordering.

Breakout Session White Paper

U.S. GOVERNMENT PRINTING OFFICE

Content Submission

The FDsys Content Submission cluster includes the processes to bring digital content into the system, and to enable Content Originators to order GPO services through FDsys. It is composed of the following functional elements:

- Types of Content
    - Deposited Content
    - Harvested Content
    - Converted Content
- Style Tools
- Content Originator Ordering

Types of Content

Digital content residing within FDsys will be composed of born digital content (deposited content), converted legacy documents (converted content), and harvested content. As the amount of deposited content submitted to GPO increases, there will be a gradual decrease in the need for legacy conversion and harvesting.

Deposited Content

Deposited content is defined as content intentionally submitted to GPO by Content Originators (e.g. Federal agency Publishers). Deposited content will include the digital object received from the Content Originator as well as corresponding customer processing requirements and additional metadata. GPO will establish best practices, templates, and standards for capturing deposited content, including metadata to capture all the customers' requirements. The system must accept all common formats used in the industry. Automatic composition tools may be applied to the digital content upon receipt to create a compliant format for the system.

The system must be able to accept deposited content that is furnished in a wide variety of formats and media, with GPO being able to convert the content into a form that is compliant with best practices. The system must also capture all agency processing information (e.g., billing information, agency reference information, etc.). Required metadata will include technical contact information (e.g., GPO 952) that relates to the agency customer requirements for rendering formats, presentation, quantity, and size, etc.

Harvested Content

Harvested content is content within the scope of dissemination programs that is gathered from Federal agency Web sites. Discovery, assessment, and harvesting tools will be used to create a content package that can be managed by FDsys.

The harvester will consist of discovery, assessment, and harvesting tools. Discovery tools will locate electronic content from Federal agency Web sites and provide information to the assessment tool. The assessment tool determines if the discovered content is within the scope of GPO dissemination programs and whether other versions of the content already exist in the system and establishes appropriate relationships between versions. The harvesting tool gathers content and available metadata.

Converted Content

Converted content is digital content created from a tangible product. GPO will continue to work with various user communities on digitizing a comprehensive collection of legacy materials. This digital collection will be made available in the public domain for permanent public access through GPO's dissemination programs.

Other forms of digitization exist currently and could possibly evolve in the future. There may be instances in which a successful conversion and/or OCR for a given tangible legacy document becomes improbable or impossible due its physical condition or characteristics. In these cases, it may be desirable to manually recreate these documents (e.g. using manual text encoding).

Style Tools

Style Tools will enable Federal Agency publishers to prepare content prior to ingest by GPO. The goal of Style Tools is to move GPO upstream in the content origination process. Style Tools must facilitate content capture, composition, collaboration, and approval, including the following functionalities:

- Accepts and validates raw or structured information.
- Composes submitted content into appropriate layouts based on user requirements.
- Creation of the best possible layout for the content.
- Allows management of work-in-progress content.
- Provides approval processes for content based on access rights and permissions.

Style Tools will work independently or in concert with existing publishing processes and will be optimized to provide content to GPO which conforms to best practices. GPO will establish and maintain a set of best practices for the creation and submission of Content Packages. The set of best practices will be loosely modeled on the traditional *GPO Style Manual* and GPO publication 300.6.

Content Originator Ordering

Content Originator Ordering provides a system interface for Content Originators to submit content, order and re-order content, and specify Content Delivery and other service options. Content Originator Ordering should allow users to discover the cost of content delivery, choose delivery options, request delivery, and discover payment/billing status for delivery of content when applicable. The system will provide the capability to create, acquire, and store metadata elements specific to ordering functions, preservation needs, version, and job specifications.

U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

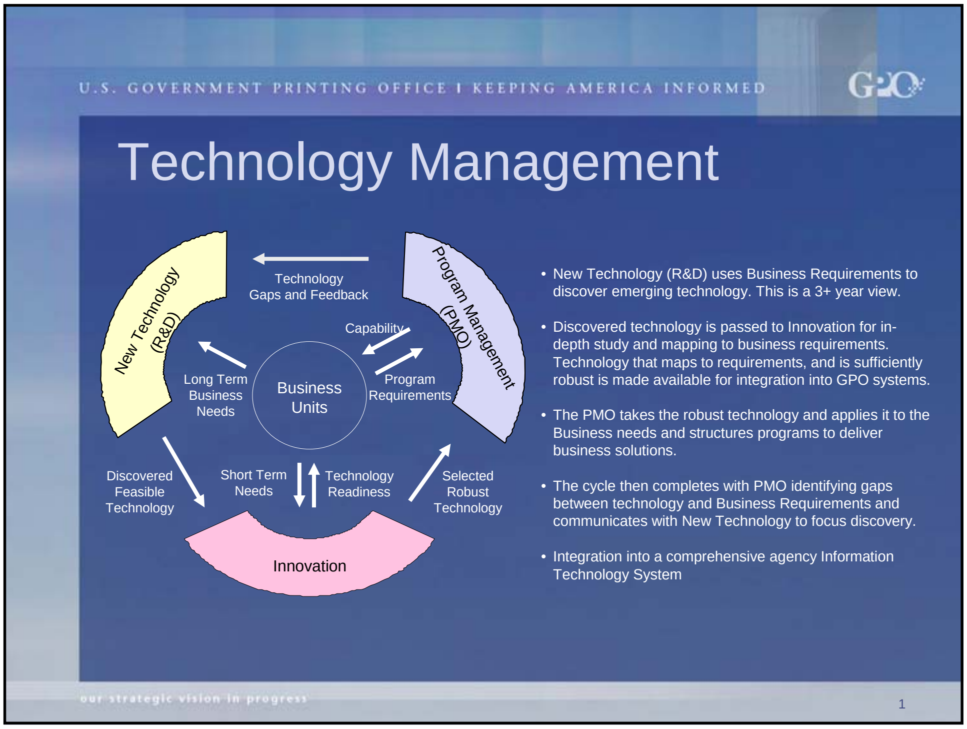
Industry Day

Content Submission

Breakout Session 3

October 6, 2005

our strategic vision in progress



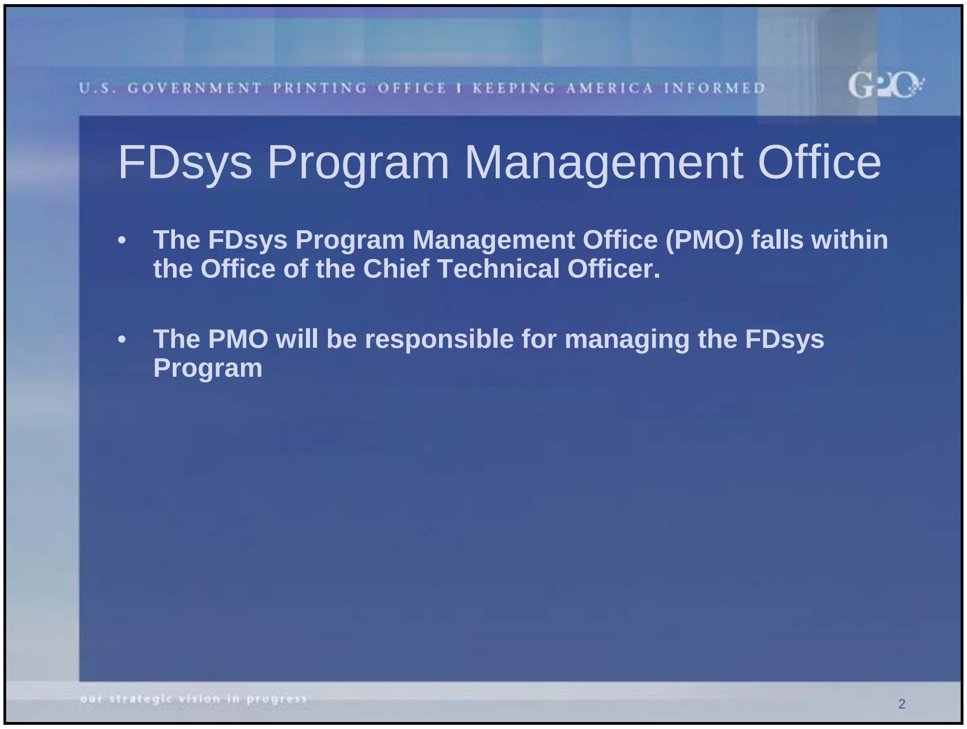
U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

Technology Management

- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System

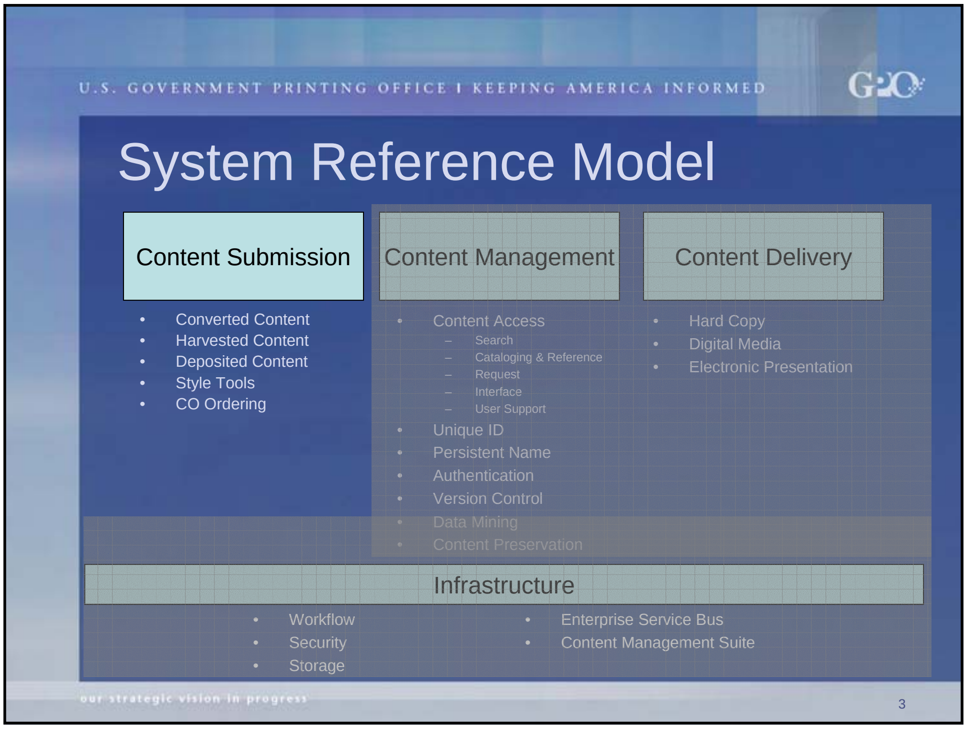
our strategic vision in progress

1



FDsys Program Management Office

- **The FDsys Program Management Office (PMO) falls within the Office of the Chief Technical Officer.**
- **The PMO will be responsible for managing the FDsys Program**



System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools
- CO Ordering

Content Management

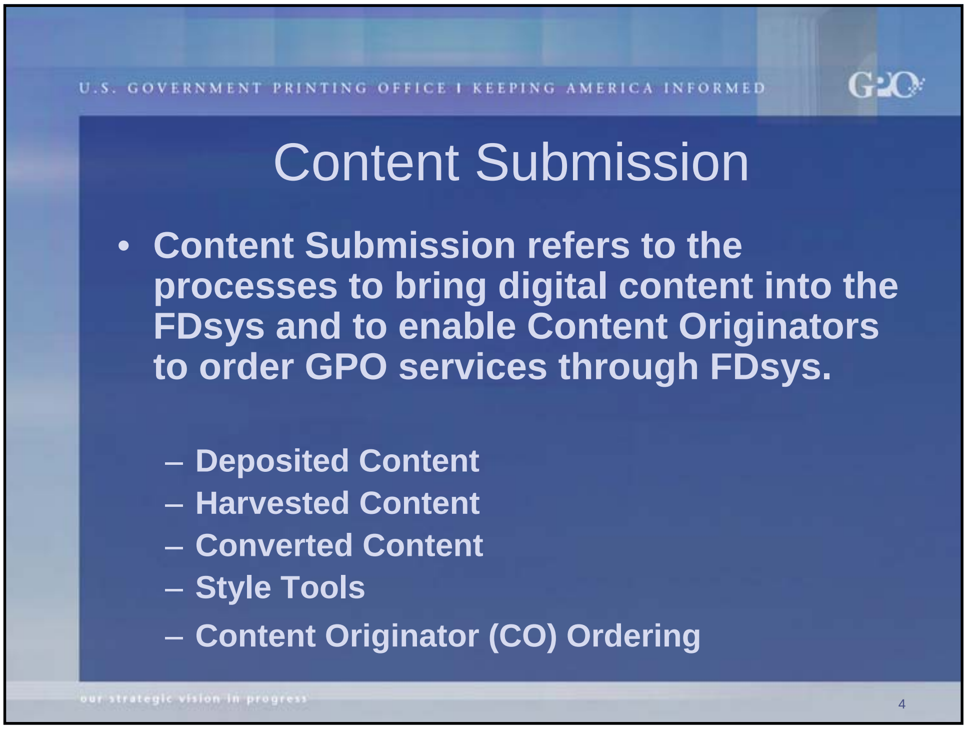
- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

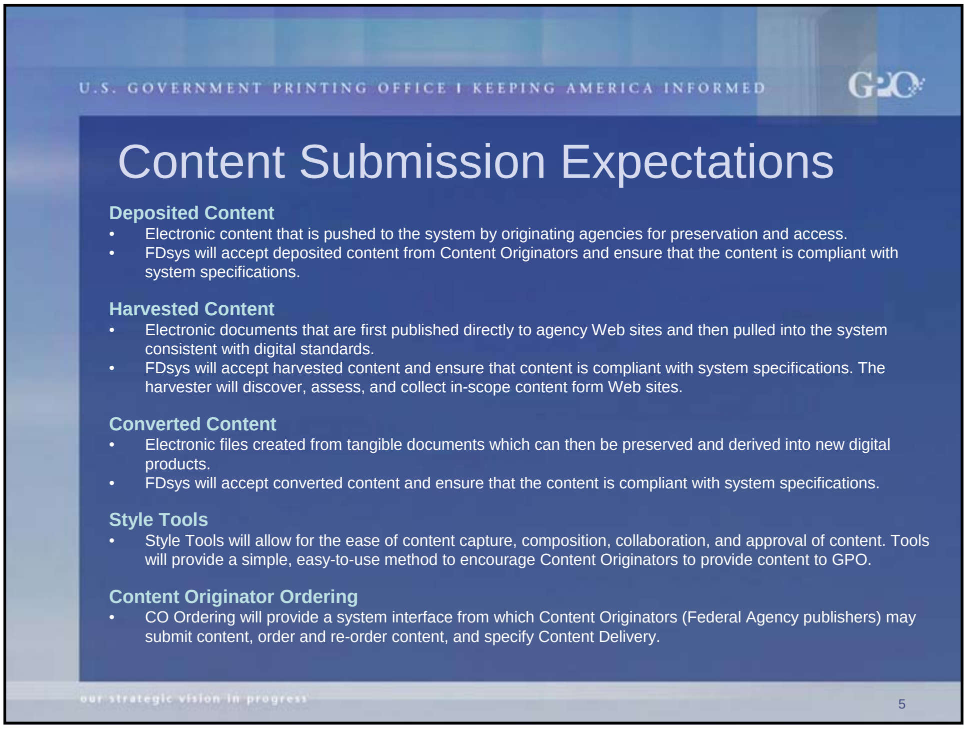
Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite



Content Submission

- **Content Submission refers to the processes to bring digital content into the FDsys and to enable Content Originators to order GPO services through FDsys.**
  - **Deposited Content**
  - **Harvested Content**
  - **Converted Content**
  - **Style Tools**
  - **Content Originator (CO) Ordering**



Content Submission Expectations

Deposited Content

- Electronic content that is pushed to the system by originating agencies for preservation and access.
- FDsys will accept deposited content from Content Originators and ensure that the content is compliant with system specifications.

Harvested Content

- Electronic documents that are first published directly to agency Web sites and then pulled into the system consistent with digital standards.
- FDsys will accept harvested content and ensure that content is compliant with system specifications. The harvester will discover, assess, and collect in-scope content form Web sites.

Converted Content

- Electronic files created from tangible documents which can then be preserved and derived into new digital products.
- FDsys will accept converted content and ensure that the content is compliant with system specifications.

Style Tools

- Style Tools will allow for the ease of content capture, composition, collaboration, and approval of content. Tools will provide a simple, easy-to-use method to encourage Content Originators to provide content to GPO.

Content Originator Ordering

- CO Ordering will provide a system interface from which Content Originators (Federal Agency publishers) may submit content, order and re-order content, and specify Content Delivery.

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 4

Infrastructure

The Infrastructure Cluster includes the system elements of Workflow, Security, Storage, Enterprise Service Bus (ESB) and Content Management Suites (CMS).

- **Workflow** is the automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules.
- **Security** includes the protection of systems (applications) against unauthorized access to or modification of information, whether in storage, processing or transit, and against the denial of service to authorized users or the provision of service to unauthorized users.
- **Storage** includes the functions associated with saving digital publications on physical media, including magnetic, optical, or other alternative technologies.
- **ESB** is an integration approach that provides a loosely-coupled, highly scalable integration infrastructure. ESB is standards-based and combines messaging, web services, data transformation, and intelligent routing to reliably connect and coordinate interactions of diverse applications across extended enterprises with transactional integrity.
- **CMS** manages information throughout its life cycle, from creation to archiving. CMS supports regulatory compliance, risk management, customer self-service, CRM applications, as well as document services, records management, forms, authoring, imaging, transformation, import/export, indexing, storage management, file systems, workflow, library services, search, collaboration, development kits, publishing, etc.

Breakout Session White Paper

U.S. GOVERNMENT PRINTING OFFICE

Infrastructure

The infrastructure of FDsys facilitates and integrates individual functional components in the system to accomplish business objectives for GPO. Infrastructure also protects the system from security breaches such as unauthorized access and unauthentic alteration to content.

Infrastructure is composed of:

- Workflow
- Security
- Storage
- Enterprise Service Bus (ESB)
- Content Management Suite (CMS)

Workflow

Workflow will be the fundamental system element in FDsys to manage business processes that are comprised of both automated and manual tasks. The types of the workflows range from fully automated processes for localized specific tasks within a functional component to system wide workflows that interact with applications and users to accomplish predefined business objectives.

The system must provide a toolset to define, execute and monitor workflows. Each workflow will be defined to support concrete business requirements, and the definition must be standards-based. The system will provide a GUI-based edit tool to create, modify and test workflow definitions before releasing workflows to production. The system shall have the capability to suspend, resume and cancel active workflows. Workflow will support monitoring current as well as history workflow activities.

Security

Security in FDsys is defined to protect the system against unauthorized access to or modification of data, whether in storage, processing or transit, and against denial of service to authorized users or the provision of service to unauthorized users. The security of the system includes the measures necessary to proactively detect, document, and counter threats. The measures and controls, including physical controls in conjunction with management, technical and procedural controls, ensure the confidentiality, integrity and availability of data processed and stored by the system.

The system must provide the capability to enforce restrictions on access to content (both authentication and authorization), to assign user rights (authorization), and to maintain system security (auditing). In addition, the system must also support the necessary technical, operational and management controls, such as monitoring capabilities for security. Security covers both internal and external system interfaces, as well as operational processes associated with FDsys.

Storage

The storage in FDsys is responsible for saving digital contents on physical media, including magnetic, optical, or other media. Following the life cycle of the content in FDsys, the storage stores content for different purposes at different stages. These include original content during content submission for submission packages, access package for user access to the content, and archival packages for preservation.

Work in Progress (WIP) Storage Deposited digital content submitted by content originators are stored in a temporary work-in-progress area , for the use of Style Tools. The storage is of temporary nature and moderate capacity, but requires fast accessibility to the content. Converted and harvested content are stored in WIP storage before content ingest as well.

Access Storage As part of the content processing, Access Content Packages (ACP) are created in FDsys . The storage for this package requires large capacity media and high performance accessibility.

Archival Storage Archival storage receives Archival Information Package (AIP) for digital content preservation. It is independent of other storage in the system and is an element of the secure repository environment for preservation processes.

Enterprise Service Bus (ESB)

FDsys is composed of many functional components (services and applications). The functional components need to be integrated or orchestrated to perform to support enterprise business processes. The system must provide an integration infrastructure to facilitate service communications and ensure interoperability between services deployed in heterogeneous hardware and software platforms, and to enable incremental implementation of FDsys.

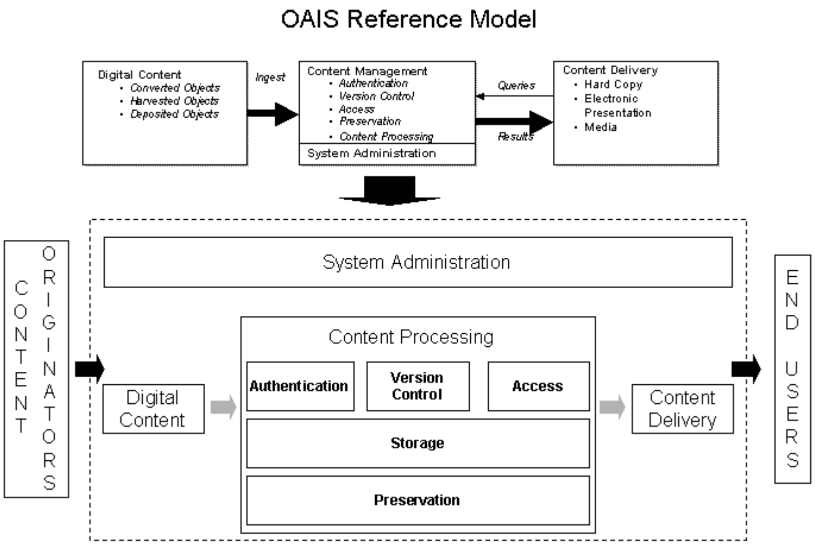
With the capability of enabling a loosely coupled, highly distributed integration infrastructure for a service-oriented architecture, the ESB is a preferred approach to fulfill the system integration requirements of FDsys. ESB infrastructure must provide the capability to deploy and remove services, internal or external, to and from the system declaratively, and to perform the XML document-based intelligent routing to destination services. ESB toolsets shall provide the capability to dynamically assemble and orchestrate enterprise business processes, as well as to monitor status of orchestrated business processes. The integration infrastructure shall also provide the capability to enable and disable selected enterprise business processes, and to cancel, suspend and resume business processes. Rule-based security control must be enforced on the administrative tasks of the ESB infrastructure management.

Content Management Suite (CMS)

Content management in FDsys manages digital content with a mandate for the preservation and dissemination of information products generated by the entire Federal government, including current, legacy and future products in all conceivable formats. Digital content must be preserved free from dependence on specific hardware and/or software.

The CMS for FDsys must manage digital information throughout its life cycle, from creation to dissemination and archiving. These functions include, but are not limited to, content creation, capture and collaboration, records management, content transformation for access, dissemination and preservation archiving, indexing, storage management, file systems, workflow, library services, access or search processing, publishing and content delivery services, integration, regulatory compliance, risk management, customer self-service, CRM applications, etc.

Conceptually, FDsys follows the Open Archival Information System reference model (OAIS), in which content and metadata are managed and stored as logical packages. A CMS for FDsys must have the capability to support multiple metadata models, and to process and preserve multiple sets of metadata for a single digital content.



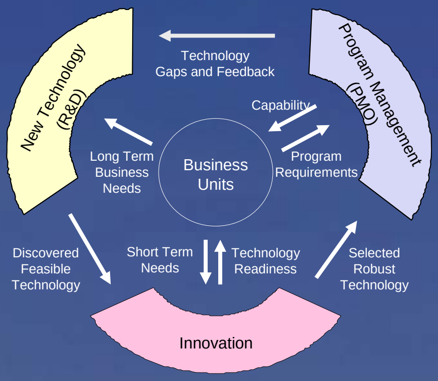
Industry Day

Infrastructure

Breakout Session 4

October 6, 2005

Technology Management



New Technology (R&D)

Technology Gaps and Feedback

Program Management (PMO)

Capability

Long Term Business Needs

Business Units

Program Requirements

Discovered Feasible Technology

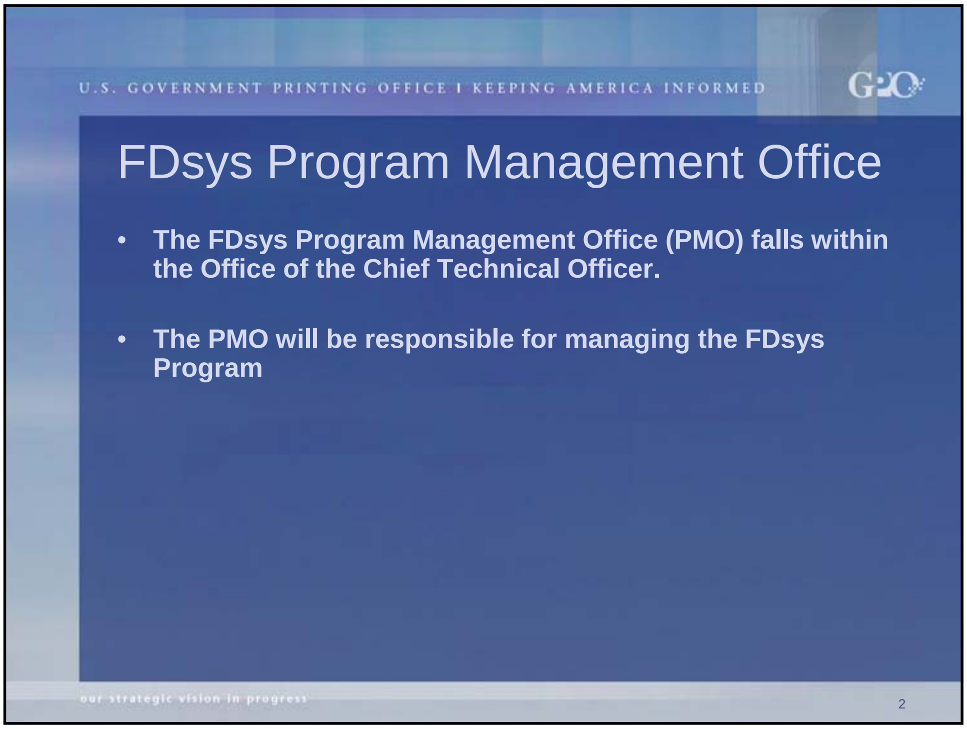
Short Term Needs

Technology Readiness

Selected Robust Technology

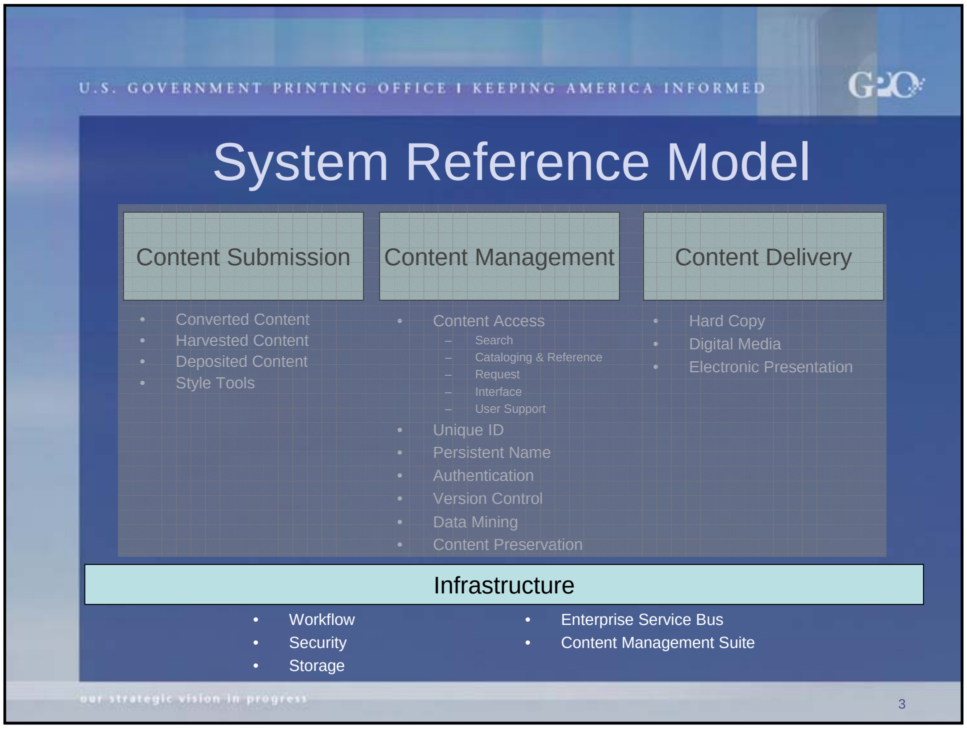
Innovation

- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System



FDsys Program Management Office

- **The FDsys Program Management Office (PMO) falls within the Office of the Chief Technical Officer.**
- **The PMO will be responsible for managing the FDsys Program**



System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

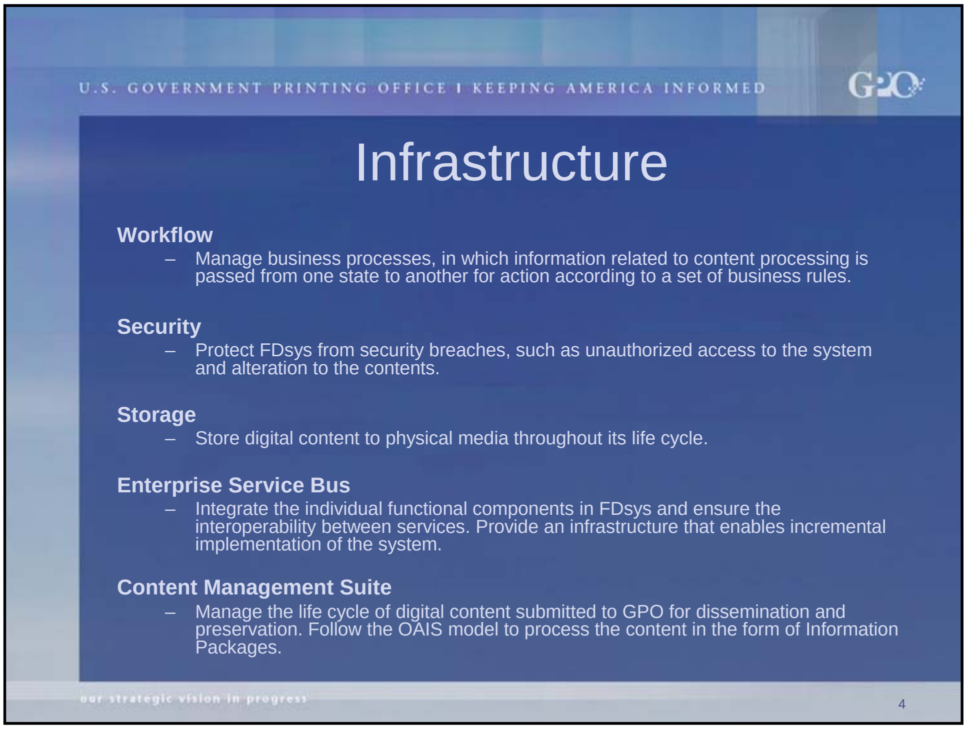
- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

Infrastructure

Workflow

– Manage business processes, in which information related to content processing is passed from one state to another for action according to a set of business rules.

Security

– Protect FDsys from security breaches, such as unauthorized access to the system and alteration to the contents.

Storage

– Store digital content to physical media throughout its life cycle.

Enterprise Service Bus

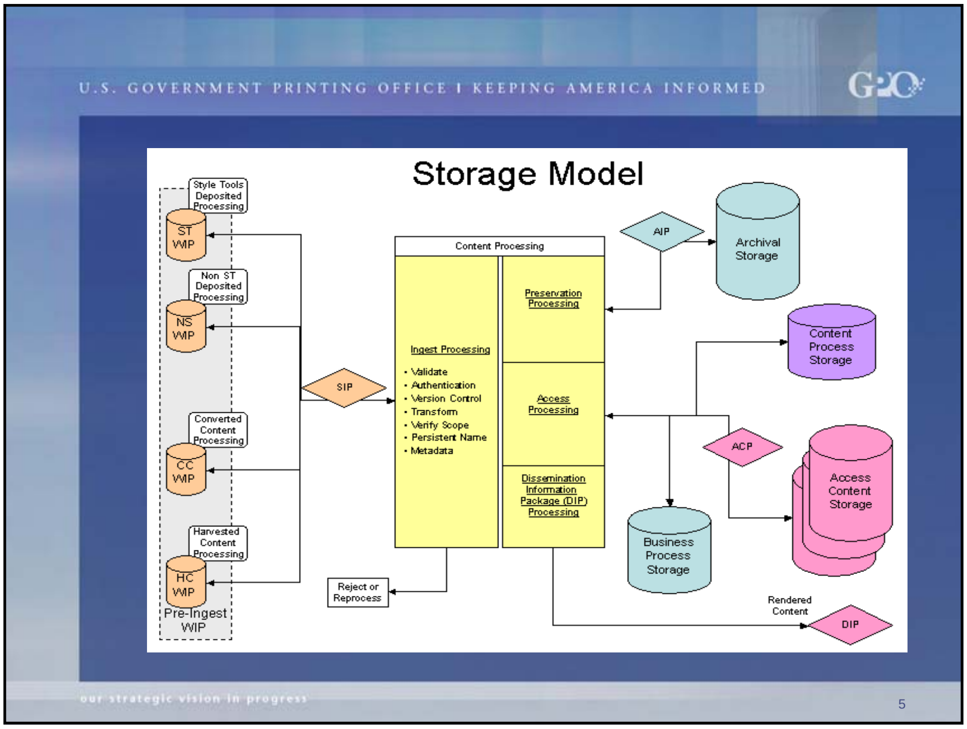
– Integrate the individual functional components in FDsys and ensure the interoperability between services. Provide an infrastructure that enables incremental implementation of the system.

Content Management Suite

– Manage the life cycle of digital content submitted to GPO for dissemination and preservation. Follow the OAIS model to process the content in the form of Information Packages.

our strategic vision in progress

4

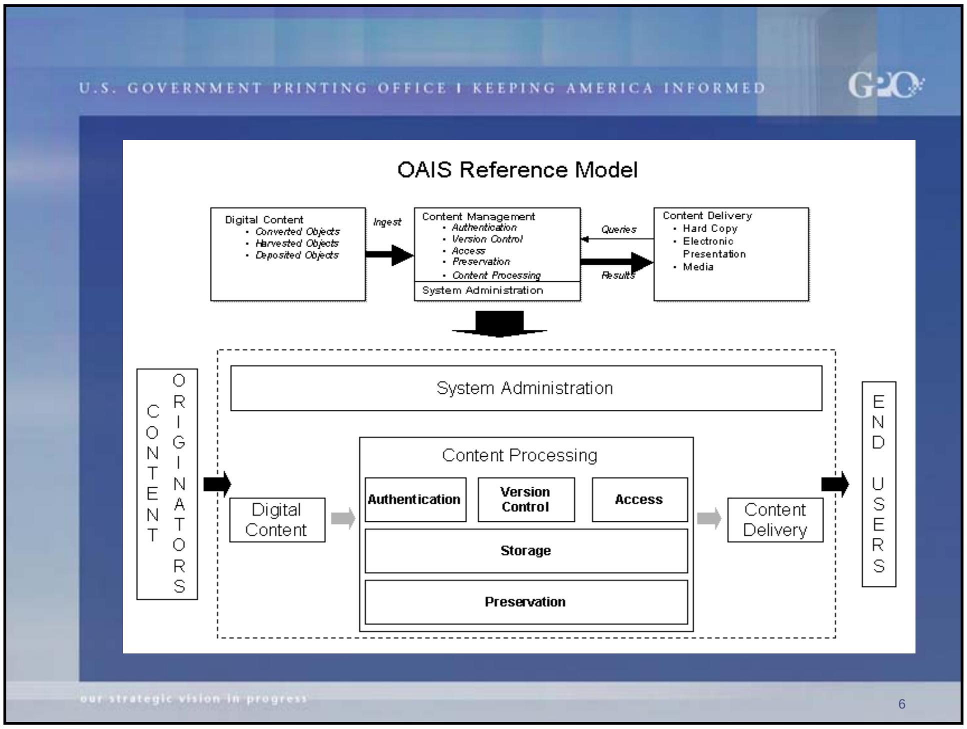


U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

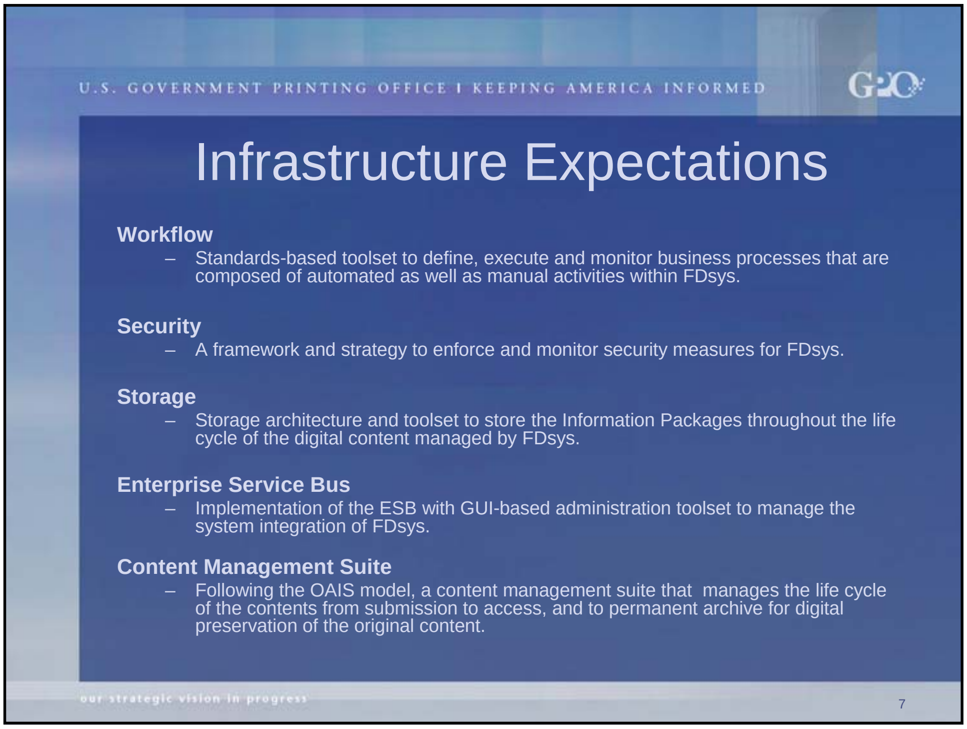
Storage Model

our strategic vision in progress

5



OAIS Reference Model



Infrastructure Expectations

Workflow

– Standards-based toolset to define, execute and monitor business processes that are composed of automated as well as manual activities within FDsys.

Security

– A framework and strategy to enforce and monitor security measures for FDsys.

Storage

– Storage architecture and toolset to store the Information Packages throughout the life cycle of the digital content managed by FDsys.

Enterprise Service Bus

– Implementation of the ESB with GUI-based administration toolset to manage the system integration of FDsys.

Content Management Suite

– Following the OAIS model, a content management suite that manages the life cycle of the contents from submission to access, and to permanent archive for digital preservation of the original content.

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 5
Content Management

The Content Management cluster includes five functions:
Unique Identifiers, Persistent Name, Authentication, Version Control, and Data Mining

- **Unique Identifiers** are character strings that uniquely identify all content within FDsys throughout the content lifecycle. The Unique ID will be assigned once, and will be indefinitely linked to the corresponding content.
- **Persistent Name** provides for the identification of each digital object independent of its location. It will provide permanence of identification, resolution of location, and be globally registered, validated, and unique.
- **Authentication** will assure customers that the information made available by GPO through the Future Digital System is authentic and/or official. There is a need for information that is reliable because it is from a trusted source, and a need to ensure the protection of data against unauthorized modification or substitution of information.
- **Version Control** will evaluate and establish the version of a piece of content and subsequently track it through its entire lifecycle, based on best practices. Version control will be called upon to analyze Content Packages and assign the appropriate version, consistent with requirements for version triggers and chain of responsibility. The chain of responsibility will be reflected in the metadata.
- **Data Mining** is associated with finding, aggregating, and associating business process information through the analysis of collections of data.

GPO U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

Breakout Session White Paper

U.S. GOVERNMENT PRINTING OFFICE

Content Management

The FDsys Content Management cluster includes the processes necessary to identify, manage, and verify digital content as it moves through the system. The FDsys Content Management cluster includes the following functional areas:

- **Unique Identifiers –** Uniquely identifying all content throughout the content lifecycle.
- **Persistent Name –** Providing permanent identification of content delivered to users.
- **Authentication –** Assuring the official/authentic nature of content.
- **Version Control –** Managing and tracking versions of content throughout the content lifecycle.
- **Data Mining –** Analyzing business process information (BPI) and metadata in order to measure performance, troubleshoot, or improve efficiency.

Unique Identifiers

Unique Identifiers are character strings that uniquely identify all content within FDsys throughout the content lifecycle. Content managed by the system will be automatically assigned a system generated alphanumeric identifier upon receipt into the system. Once assigned, a Unique Identifier cannot be reused and will be linked indefinitely to the corresponding content. The uniqueness of the assigned identifier ensures that the identifier will refer to only one object. All assigned unique identifiers will be recorded and used in metadata.

Persistent Name

Persistent Name provides for the identification of content independent of its location. Persistent Naming allows for an interoperable schema of identifiers that uniquely identify content, provides permanence of identification, resolution of location, and support access to information about the content. Persistent Names will be unique, validated, and globally registered. A resolution system will locate and provide access to content and metadata associated with assigned persistent names.

Persistent Names will be assigned to content packages when ingested into the system. Once assigned, a persistent name will not be reused within the system. Each name will be human readable and interoperable with legacy schemes and other systems. Persistent Name information will be recorded in metadata and all Persistent Name transactions will be logged.

Authentication

Authentication will assure customers that the information made available by GPO through the Future Digital System is authentic and/or official. There is a need for information that is reliable because it is from a trusted source, and a need to certify that content is free of unauthorized modification or substitution of information. The system will be capable of certifying content as authentic and/or official at both the document and granular levels. Authentication will verify and validate deposited, harvested, and converted content. Authentication will be able to notify when, where, by whom, and what changes were made to the content. Authentication will be conveyed by users through the use of integrity marks.

GPO currently has an operational Public Key Infrastructure (PKI), for authentication of content. The FDsys will be able to interface with the PKI wherever feasible to enable end users to determine that files are unchanged since they were authenticated by GPO and help establish a clear chain of custody for electronic documents.

Version Control

Version Control will manage the version of a piece of content and subsequently track it through its entire life cycle. A version is a unique manifestation of a publication. The system will be able to determine whether a change to a publication constitutes a new version by referring to that document's version trigger. The version trigger contains a tolerance level for changes to a publication. If the changes to a document surpass its threshold in the version trigger, a new version is created and labeled by Version Control. The system will be able to assign and store version triggers in the individual metadata for each publication. The metadata managed by Version Control will also reflect the chain of responsibility for the publication.

Data Mining

Data Mining consists of the tools and processes for finding, analyzing, and presenting business process information (BPI) in order to measure performance, troubleshoot, or improve efficiency. Data Mining includes the logging of all system processes and transactions. Subject to privacy and legal restrictions, Data Mining will be able to extract data from the system's entire BPI and metadata population, including selected access to external data repositories, on a permission basis. Data Mining will include a robust analysis toolset, including statistical tools and the ability to incorporate taxonomies and ontologies. Users will be able to create and save custom reports.

U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

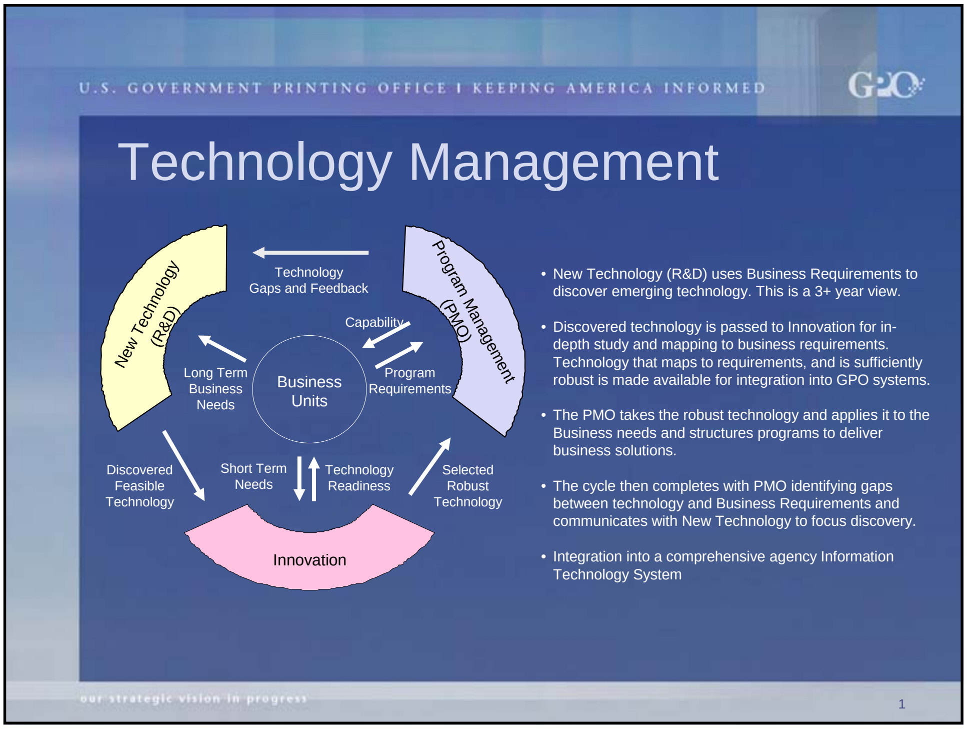
Industry Day

Content Management

Breakout Session 5

October 6, 2005

our strategic vision in progress

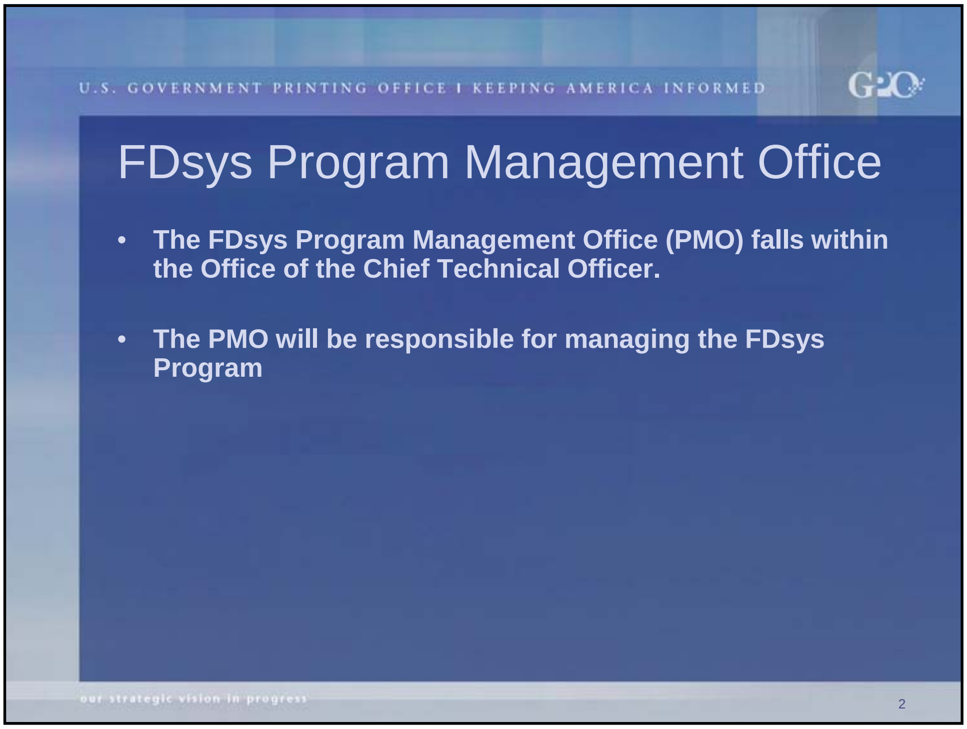


U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

Technology Management

- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System

our strategic vision in progress



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

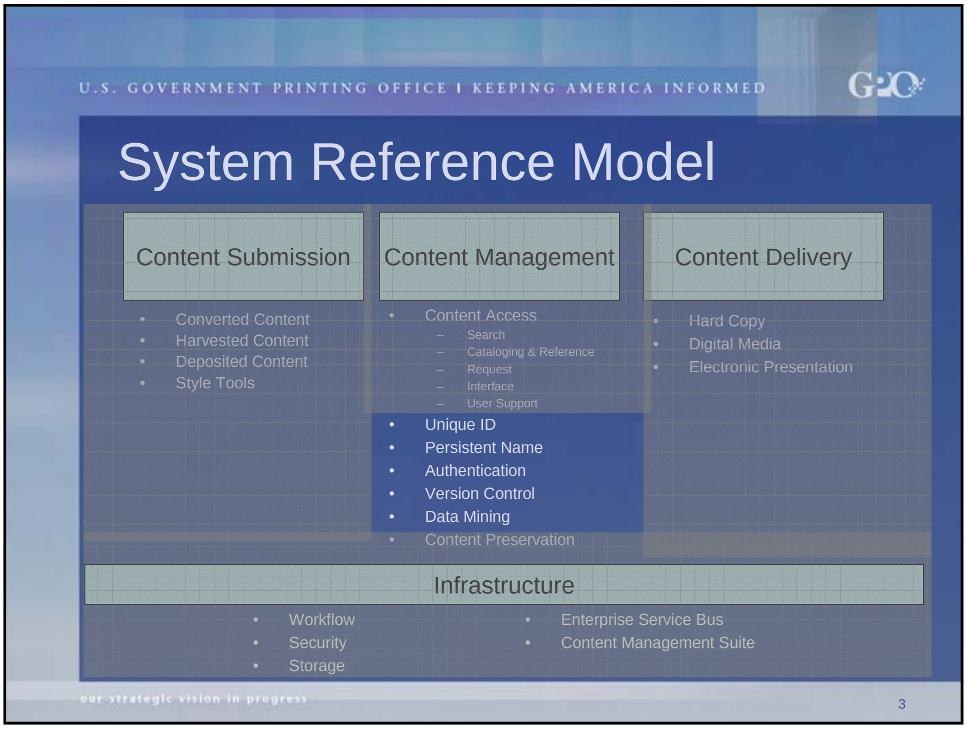
GPO

FDsys Program Management Office

- The FDsys Program Management Office (PMO) falls within the Office of the Chief Technical Officer.
- The PMO will be responsible for managing the FDsys Program

our strategic vision in progress

2



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

GPO

System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

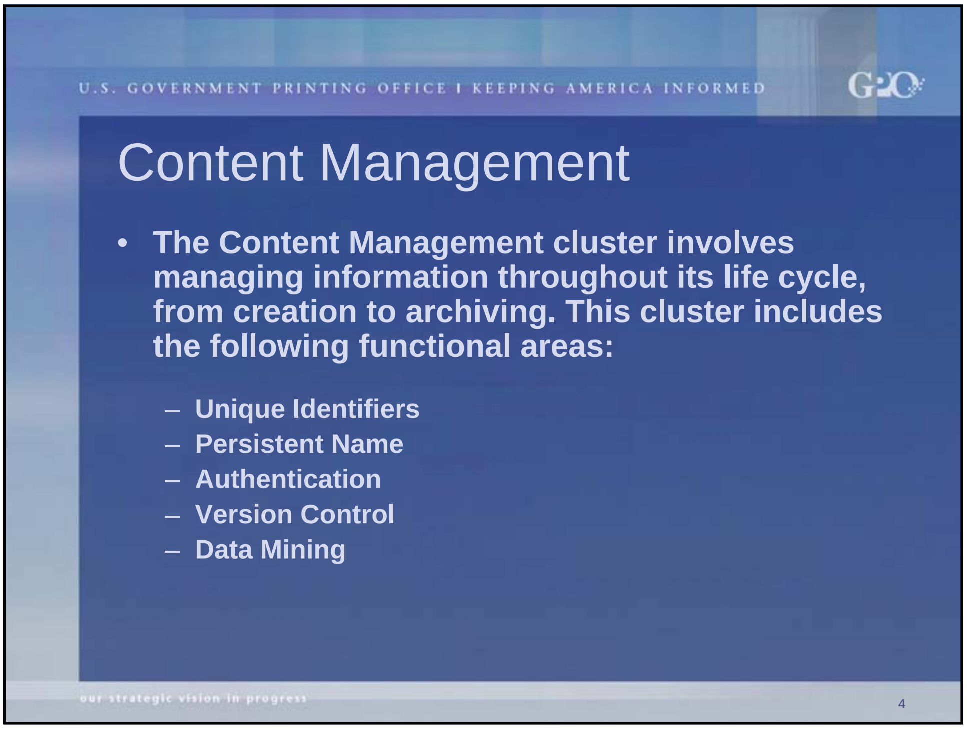
- Hard Copy
- Digital Media
- Electronic Presentation

Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite

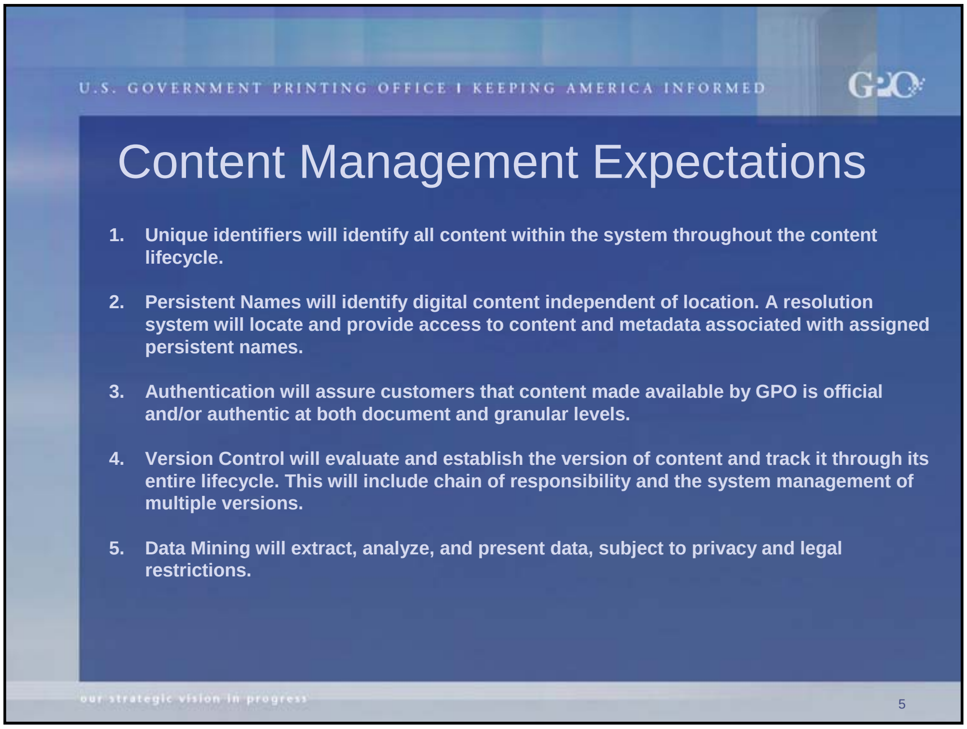
our strategic vision in progress

3



Content Management

- The Content Management cluster involves managing information throughout its life cycle, from creation to archiving. This cluster includes the following functional areas:
  - Unique Identifiers
  - Persistent Name
  - Authentication
  - Version Control
  - Data Mining



Content Management Expectations

1. Unique identifiers will identify all content within the system throughout the content lifecycle.
2. Persistent Names will identify digital content independent of location. A resolution system will locate and provide access to content and metadata associated with assigned persistent names.
3. Authentication will assure customers that content made available by GPO is official and/or authentic at both document and granular levels.
4. Version Control will evaluate and establish the version of content and track it through its entire lifecycle. This will include chain of responsibility and the system management of multiple versions.
5. Data Mining will extract, analyze, and present data, subject to privacy and legal restrictions.

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 6

Content Delivery

The Content Delivery cluster includes mechanisms for delivering content that fits the requirements of the Content Originator or End User. The delivery methods identified are hard copy, electronic presentation, and digital media.   Content Delivery requirements reflect the expansion of electronic presentation as the primary dissemination method for Government information.

U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

Breakout Session White Paper

U.S. GOVERNMENT PRINTING OFFICE

Content Delivery

Background

Content Delivery refers to mechanisms for delivering content in a manner that fits the requirements of Content Originators or End Users. Identified methods include **hard copy, electronic presentation, and digital media**.   Content Delivery requirements reflect the expansion of electronic presentation as the primary dissemination method for Government information.

The system will retrieve content from storage based on a user request, and package the content for delivery to the user. Content prepared for dissemination may include integrity marks certifying the content as official and/or authentic. Content Delivery includes:

- Transfer of content to a Service Provider for Hard Copy output,
- Delivery to an End User in electronic format,
- Delivery to an End User via Digital Media (e.g., CDs, DVDs), and
- Delivery to Digital Media (devices external to the system such as PDAs, MP3 players, etc.).

Content may be pushed from the system to users (e.g., RSS feeds, e-mail, raw data feeds) and pulled from the system by users (e.g., FTP).

A database of Service Providers will be managed by GPO. This database will allow Service Providers to electronically maintain profiles (e.g., equipment, capabilities) and will allow authorized GPO users to manage Service Provider performance information (e.g., quality and performance data). Content Originators and users will be able to select Service Providers based on needs and requirements. Tracking of orders and logging of Service Provider activities (e.g., job received, job complete) will also be possible.

The system will comply with GPO Quality Assurance standards for content delivery, best practices and guidelines regarding usability for electronic content (e.g., The Research-Based Web Design & Usability Guidelines function as best practices for Federal Agencies), accessibility laws and regulations (e.g., Section 508 of the Rehabilitation Act Amendments of 1998), and W3C guidelines.

Hard Copy

Tangible printed content.

The system will be capable of determining the most cost effective method for hard copy output (offset press, digital printing, etc.), and will deliver an appropriate content package to the Service Provider. Content packages suitable for print on demand output (hard copy produced in a short production cycle time and typically in small quantities) will also be generated by the system. Hard copy output may be requested either at the time of content submission, or as an access order.

Electronic Presentation

Dynamic and temporary representation of content in digital format; strongly dependent upon file format and user's presentation device.

Electronic presentation will be the primary dissemination method for authentic and/or official Government content. The system will deliver electronic formats for multiple computer platforms (e.g., Windows, Macintosh, UNIX) and non-desktop electronic devices (e.g., PDAs, MP3 players, e-books). Multiple formats will be available for delivery to End Users.

Digital Media

Intermediary mechanism consisting of data storage devices to deliver content to users' storage or display devices.

Digital media includes the following:

- Data storage devices (CD, DVD, etc.)
- Wireless handheld devices (PDAs, MP3 players, e-books, etc.)
- Future media (flexible electronic displays, etc.)
- Storage at user sites

The system will package content for transfer via digital media (e.g., CDs, DVDs) and transfer to digital media (devices such as PDAs, etc.). Content may be pushed to the user's device or pulled from the system.

U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

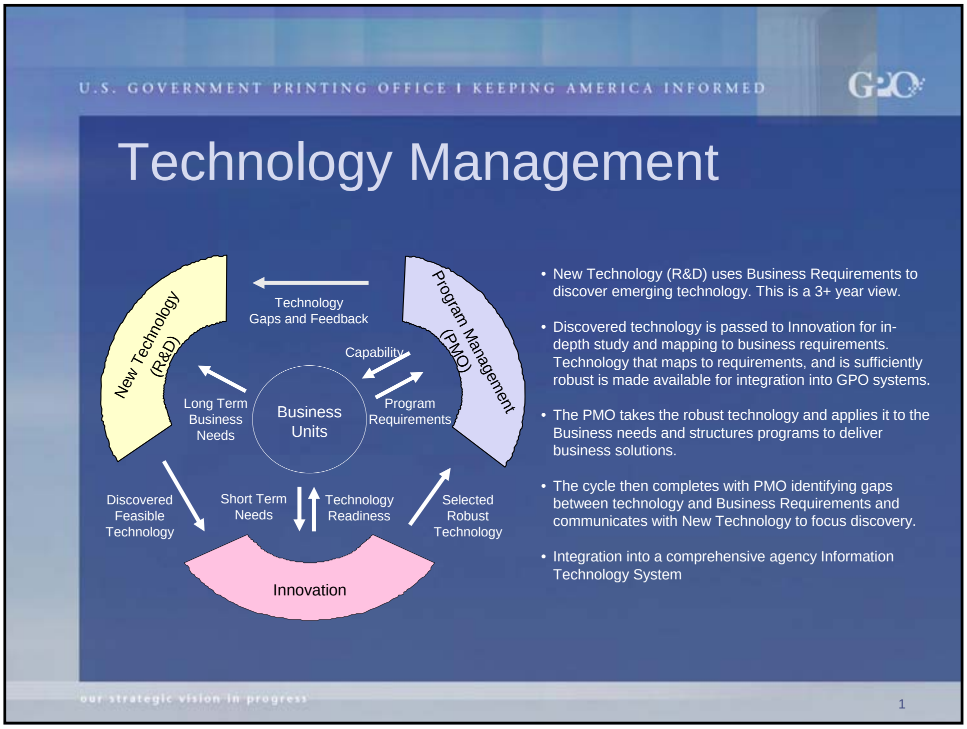
Industry Day

Content Delivery

Breakout Session 6

October 6, 2005

our strategic vision in progress

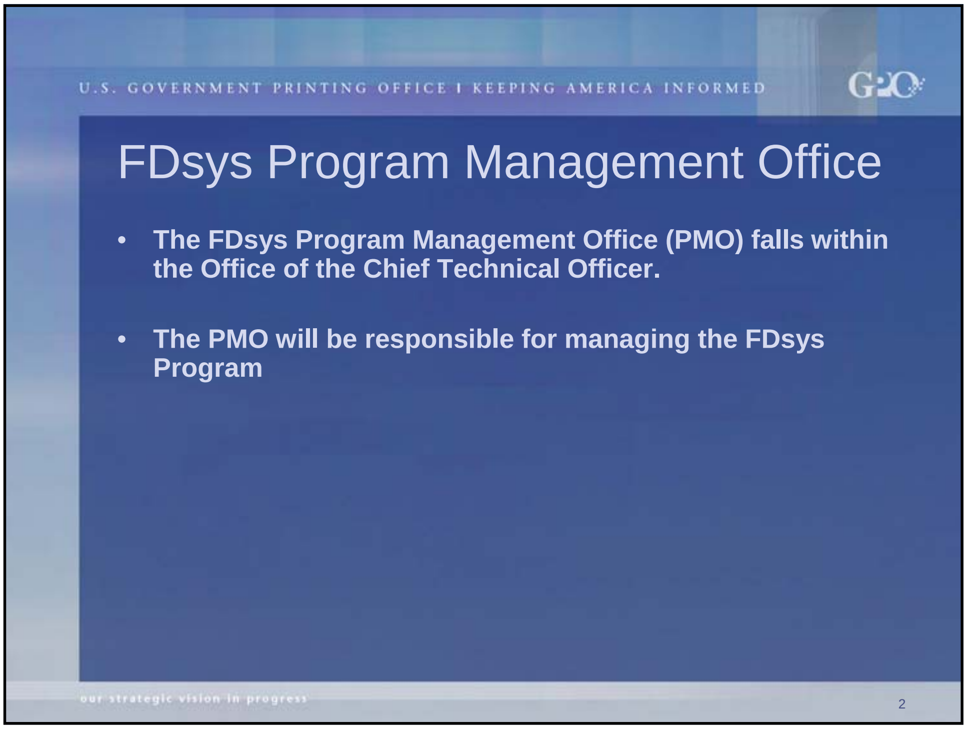


U.S. GOVERNMENT PRINTING OFFICE I KEEPING AMERICA INFORMED

Technology Management

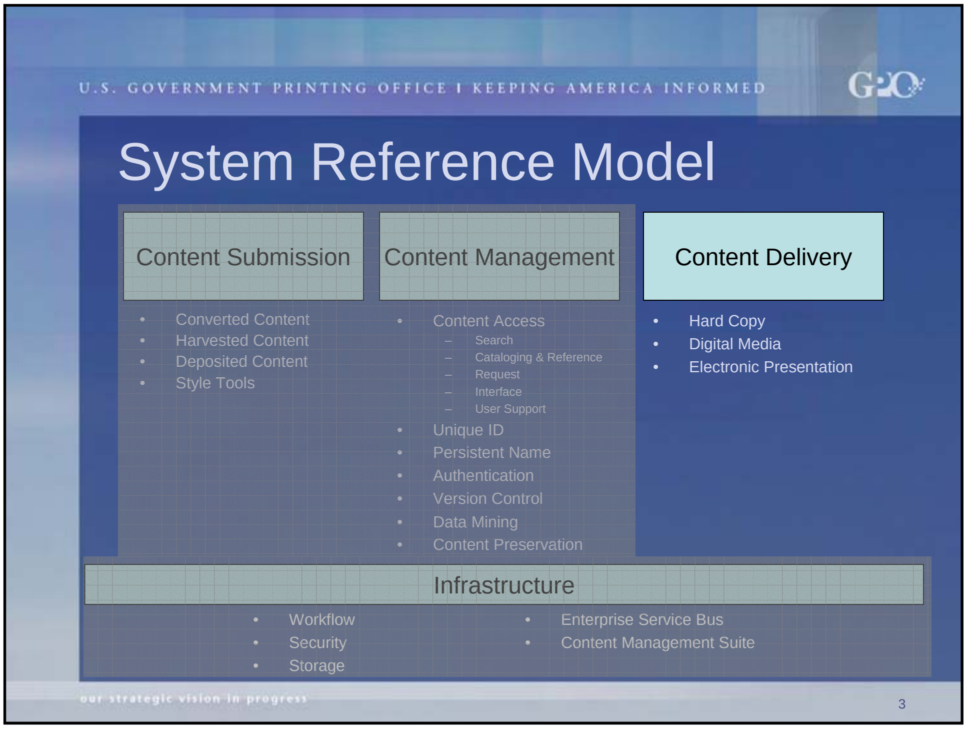
- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System

our strategic vision in progress



FDsys Program Management Office

- **The FDsys Program Management Office (PMO) falls within the Office of the Chief Technical Officer.**
- **The PMO will be responsible for managing the FDsys Program**



System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

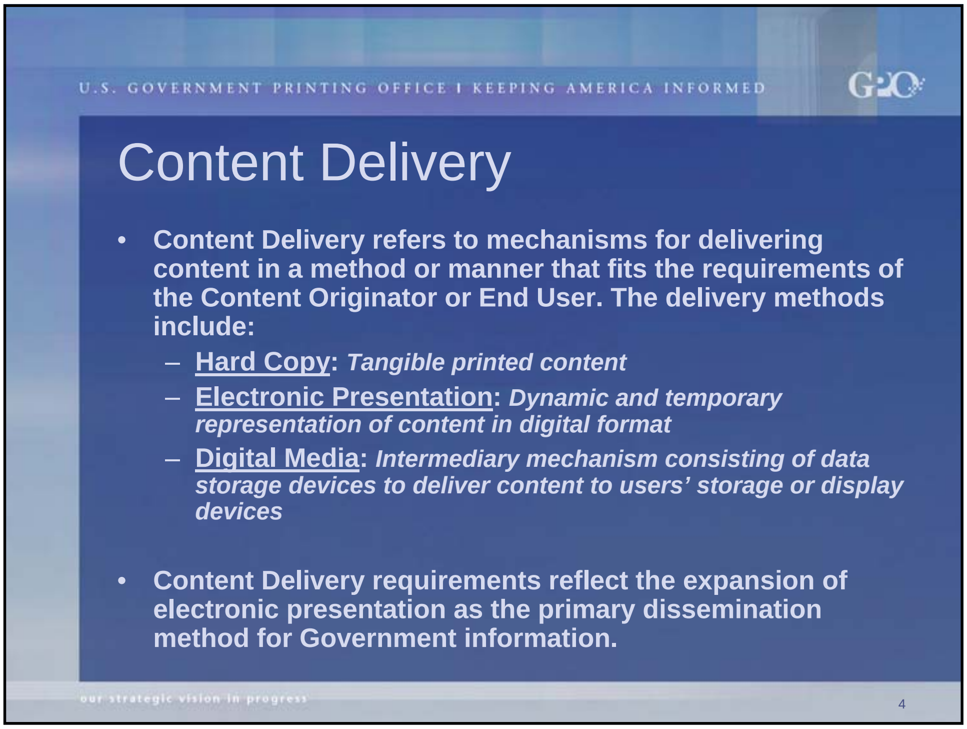
- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

Infrastructure

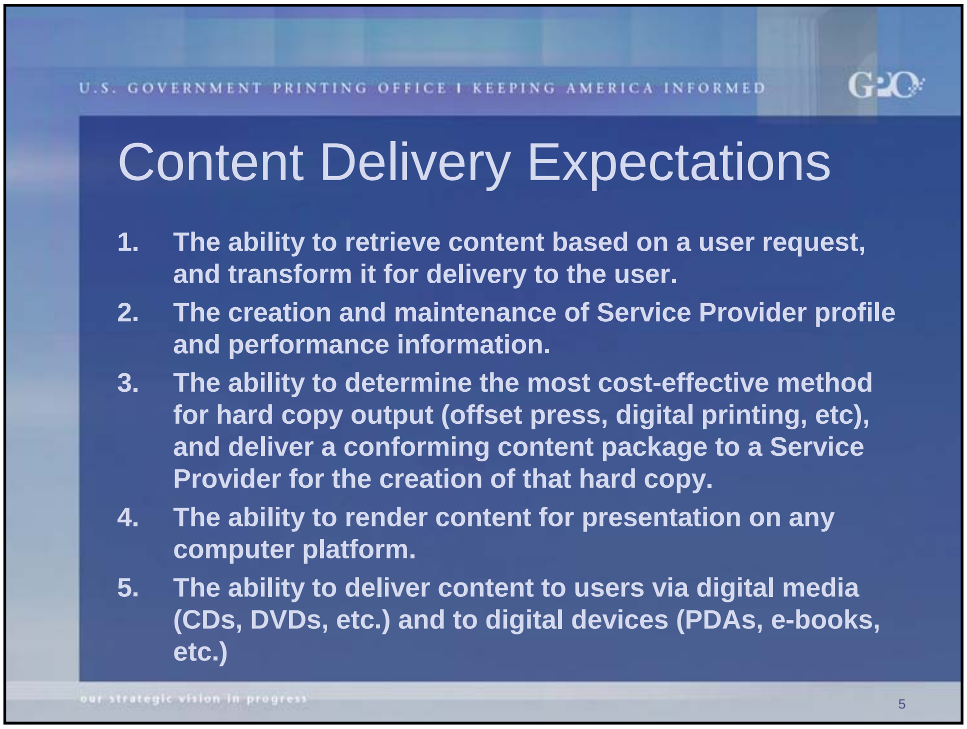
- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite



Content Delivery

- Content Delivery refers to mechanisms for delivering content in a method or manner that fits the requirements of the Content Originator or End User. The delivery methods include:
  - Hard Copy: *Tangible printed content*
  - Electronic Presentation: *Dynamic and temporary representation of content in digital format*
  - Digital Media: *Intermediary mechanism consisting of data storage devices to deliver content to users' storage or display devices*

- Content Delivery requirements reflect the expansion of electronic presentation as the primary dissemination method for Government information.



Content Delivery Expectations

1. The ability to retrieve content based on a user request, and transform it for delivery to the user.
2. The creation and maintenance of Service Provider profile and performance information.
3. The ability to determine the most cost-effective method for hard copy output (offset press, digital printing, etc), and deliver a conforming content package to a Service Provider for the creation of that hard copy.
4. The ability to render content for presentation on any computer platform.
5. The ability to deliver content to users via digital media (CDs, DVDs, etc.) and to digital devices (PDAs, e-books, etc.)

The FDsys Industry Day

A U.S. Government Printing Office Event

October 6, 2005

Breakout 7
Content Access

The Content Access cluster will be the primary interface between End Users and the FDsys.

Content Access includes the following:

- **Search -** Performing queries on content and metadata so that content can be retrieved from storage and delivered to users.
- **Cataloging -** Creating descriptive metadata that conform to accepted standards, and support access and delivery of standard bibliographic records.
- **Reference Tools -** Creating lists and resources that assist users in locating and accessing content.
- **Request -** Processing no-fee and fee based content delivery requests.
- **Interface -** Developing and managing user interactions with the system.
- **User Support -** Providing answers to user questions and directing users to content and services.

U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

Breakout Session White Paper

U.S. GOVERNMENT PRINTING OFFICE

Content Access

Content Access will be the primary interface between End Users and the FDsys. Content Access includes the following:

- **Search** – Performing queries on content and metadata so that content can be retrieved from storage and delivered to users.
- **Cataloging** – Creating descriptive metadata that conform to accepted standards, and support access and delivery of standard bibliographic records.
- **Reference Tools -** Creating lists and resources that assist users in locating and accessing content.
- **Request** – Processing no-fee and fee based content delivery requests.
- **Interface** – Developing and managing user interactions with the system.
- **User Support** – Providing answers to user questions and directing users to content and services.

Search

Search executes queries on content and metadata so that content can be retrieved from storage and delivered to users. FDsys Search tools should meet or exceed industry standards for search and retrieval technology. More than one Search tool may be used to meet the needs of all User Classes who will be searching the system. The FDsys Search tools must handle user searches of metadata and content both simultaneously and separately across multiple internal and external storage levels. Search must have the ability to search multiple media, file formats, and levels of granularity. Search should produce a highly relevant, organized, usable, and detailed results list that provides the location and description of content. Search tools should provide innovative methods for users to access information related to their query. Search must include accessible and customizable graphical user interfaces that allow all users to submit and refine queries, filter results, and export results sets.

Cataloging

Cataloging tools create descriptive metadata that conform to accepted standards and support access and delivery of standard bibliographic records. Content packages in scope for cataloging are the final published versions of authentic U.S. Government publications. The Cataloging process will use applicable descriptive metadata elements, including metadata created by the publishing agency for harvested documents. GPO will also acquire bibliographic metadata from external Content Originators and Service Providers. GPO will provide metadata records to various users in a variety of standard formats such as MARC and ONIX.

Reference Tools

Reference tools are finding aids, bibliographies, and other resources that assist users in locating and accessing content. Reference tools will have the ability to create, acquire, and store metadata, and link to metadata or content. Lists, in the context of Reference tools, may be static pages produced from report generation capabilities or dynamic results lists from searches. These searches may be pre-configured or individually created for one-time use. Reference tools will also take advantage of emerging semantic or clustering technologies, or other developments in the field of knowledge management.

Request

Request will allow End Users to request delivery of content packages available from FDsys. Request must have the capability to process no-fee and fee-based content delivery requests. An example of a no-fee Request for delivery is downloading a PDF document that is within scope of the Federal Depository Library Program. An example of a fee-based Request for delivery is using a shopping cart function to order a publication from an eCommerce Web site. For fee-based content, Request must provide the capability for End Users discover the cost of content delivery, choose delivery options, and submit payment for delivery. Request must ensure that customer transactions can be conducted in a secure environment. Request will have the ability to interact with GPO's IT infrastructure for a variety of services, including financial and inventory control systems. Request must enable customers to securely store and access order histories, user preferences for delivery options, and preferred payment methods.

Interface

The Interface functional area will allow users to develop and manage user interactions with the system. Graphical user interfaces are a key component of this functional area and will consist of a workbench, which is a set of user tools related to performing a system function. GPO will develop interfaces for internal and external user classes allowing users to perform authorized functions. Users will have the ability to opt-in to acquire the capability to customize interfaces in order to create an environment better suited for their needs and preferences. The system should provide the capability to integrate Search, Reference Tools, Request, and User Support seamlessly into an End User interface. A public End User interface will be provided to allow users to access official Federal Government information without submitting personal information.

Support

GPO has a strong commitment to provide superior customer service and user support. This commitment spans from assisting Content Originators at the stage of content creation to providing services that assist users in using GPO's diverse array of tangible and electronic products. User Support will provide answers to user questions and direct them to content and services. These services may or may not be delivered in conjunction with Content Delivery. User Support services include a helpdesk and knowledge base, interactive training, real-time alert services, and services that provide the capability for users to receive personalized support based on their stored preferences. User support will also be provided in conjunction with the public End User interface and will not require users to submit personal information to the system.

U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

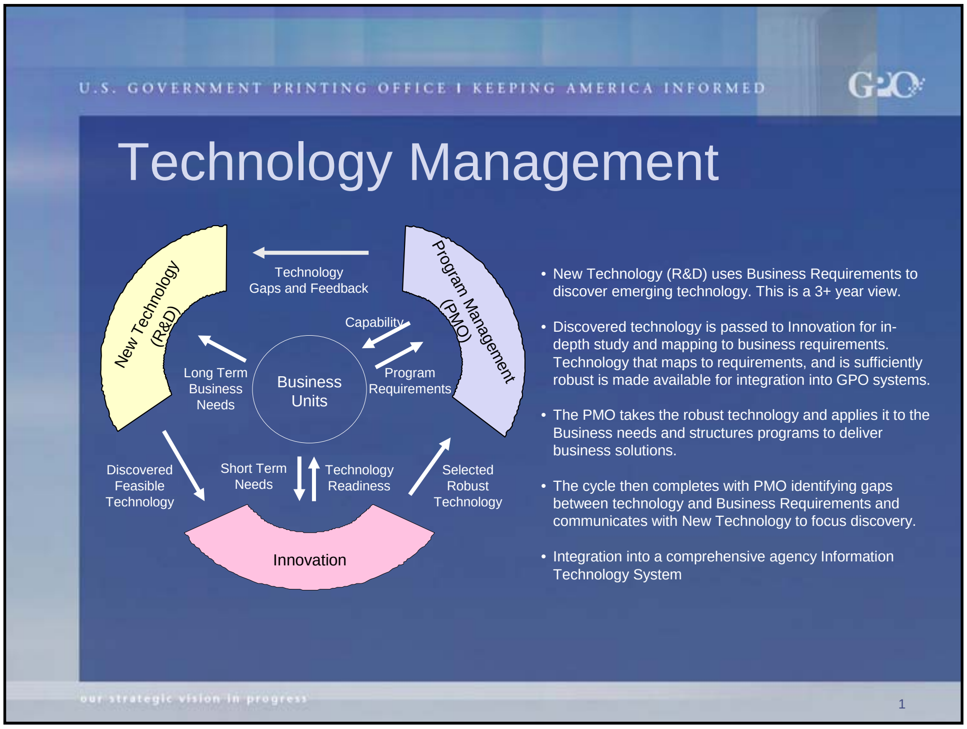
Industry Day

Content Access

Breakout Session 7

October 6, 2005

our strategic vision in progress



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

Technology Management

New Technology (R&D)

Technology Gaps and Feedback

Program Management (PMO)

Capability

Long Term Business Needs

Business Units

Program Requirements

Discovered Feasible Technology

Short Term Needs

Technology Readiness

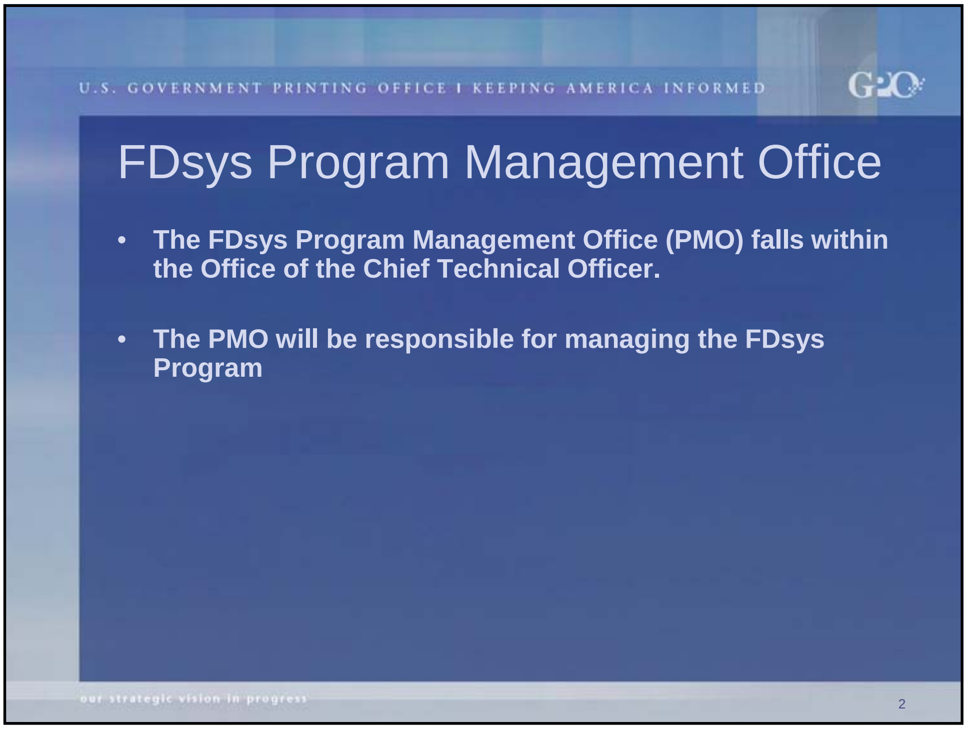
Selected Robust Technology

Innovation

- New Technology (R&D) uses Business Requirements to discover emerging technology. This is a 3+ year view.
- Discovered technology is passed to Innovation for in-depth study and mapping to business requirements. Technology that maps to requirements, and is sufficiently robust is made available for integration into GPO systems.
- The PMO takes the robust technology and applies it to the Business needs and structures programs to deliver business solutions.
- The cycle then completes with PMO identifying gaps between technology and Business Requirements and communicates with New Technology to focus discovery.
- Integration into a comprehensive agency Information Technology System

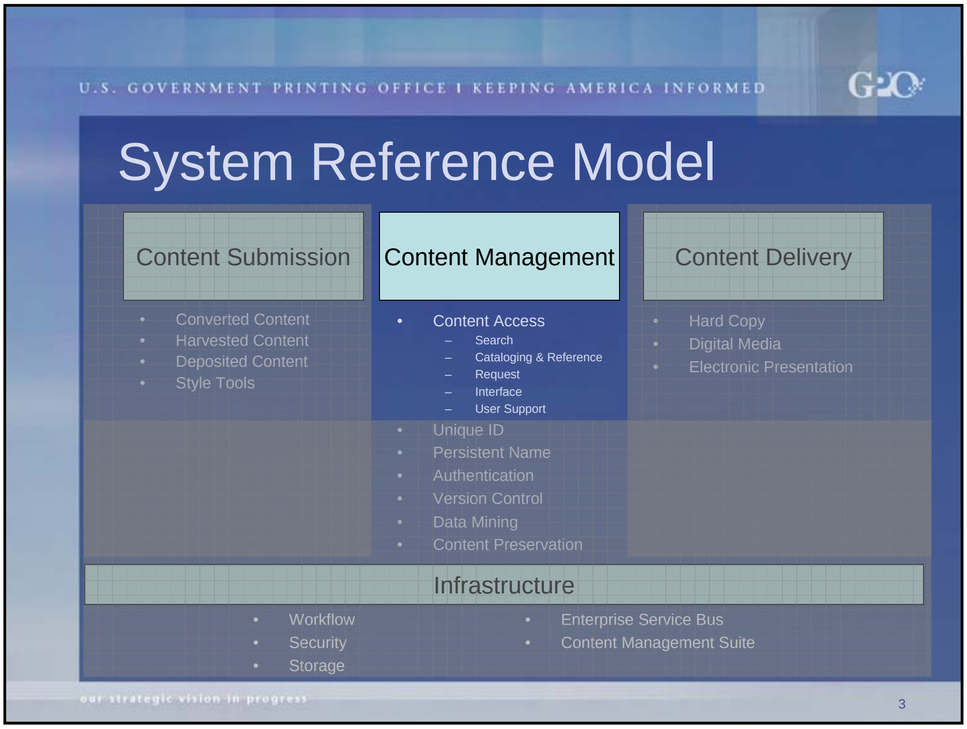
our strategic vision in progress

1



FDsys Program Management Office

- **The FDsys Program Management Office (PMO) falls within the Office of the Chief Technical Officer.**
- **The PMO will be responsible for managing the FDsys Program**



System Reference Model

Content Submission

- Converted Content
- Harvested Content
- Deposited Content
- Style Tools

Content Management

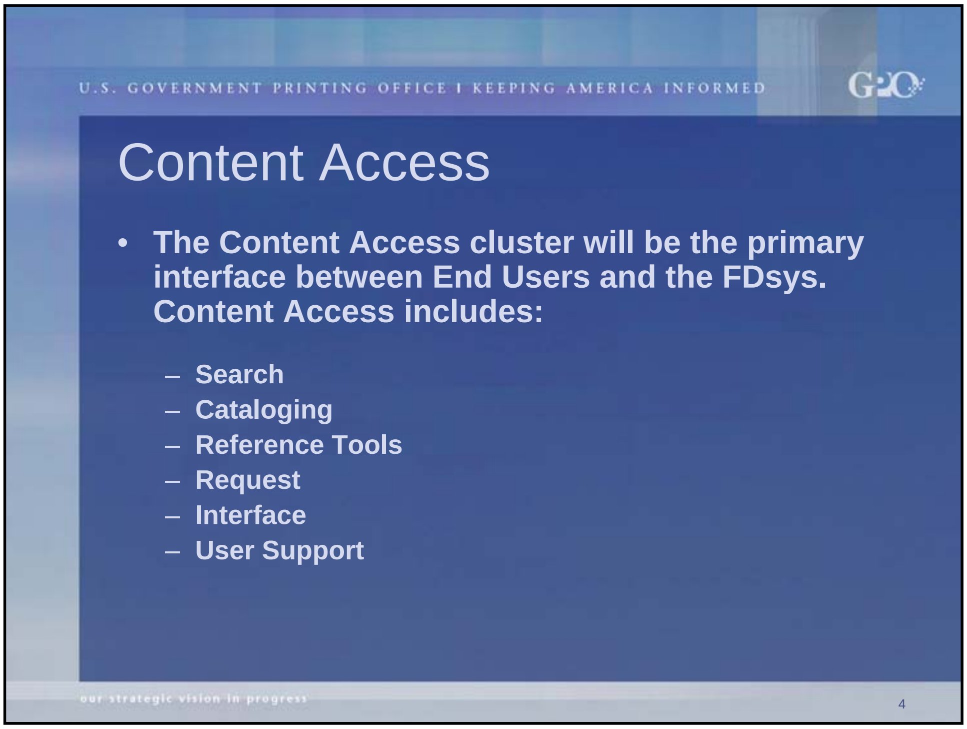
- Content Access
  - Search
  - Cataloging & Reference
  - Request
  - Interface
  - User Support
- Unique ID
- Persistent Name
- Authentication
- Version Control
- Data Mining
- Content Preservation

Content Delivery

- Hard Copy
- Digital Media
- Electronic Presentation

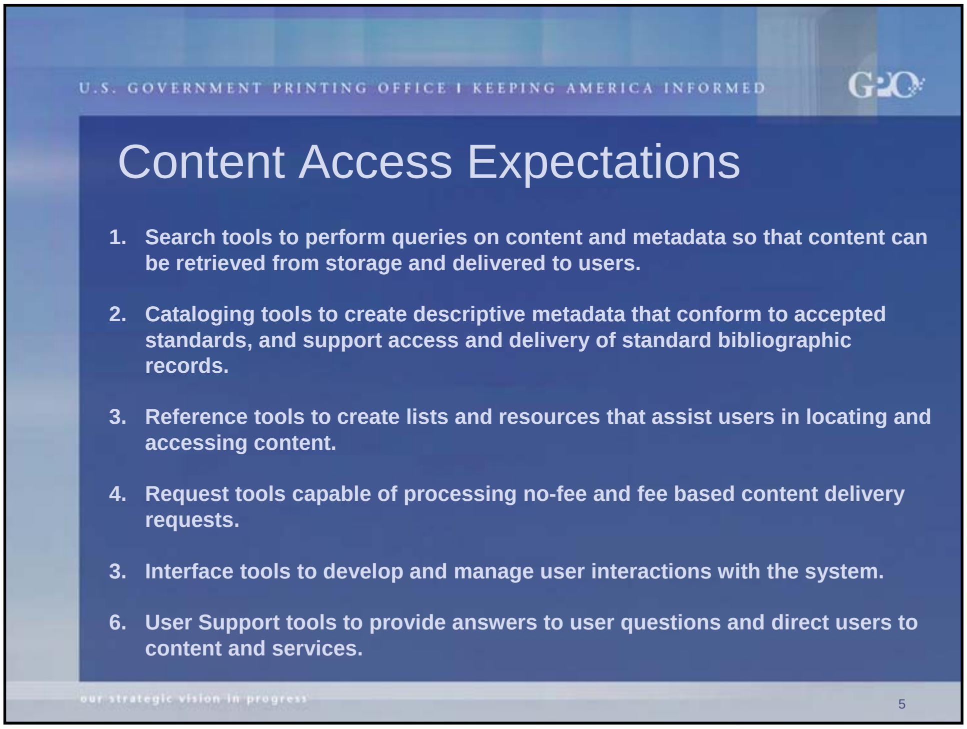
Infrastructure

- Workflow
- Security
- Storage
- Enterprise Service Bus
- Content Management Suite



Content Access

- **The Content Access cluster will be the primary interface between End Users and the FDsys. Content Access includes:**
  - **Search**
  - **Cataloging**
  - **Reference Tools**
  - **Request**
  - **Interface**
  - **User Support**



Content Access Expectations

1. Search tools to perform queries on content and metadata so that content can be retrieved from storage and delivered to users.

2. Cataloging tools to create descriptive metadata that conform to accepted standards, and support access and delivery of standard bibliographic records.

3. Reference tools to create lists and resources that assist users in locating and accessing content.

4. Request tools capable of processing no-fee and fee based content delivery requests.

3. Interface tools to develop and manage user interactions with the system.

6. User Support tools to provide answers to user questions and direct users to content and services.

GPO's Future Digital System:

System Releases and Capabilities

Version 2.0

GPO FDsys Reference Document

September 30, 2006

Document Change/Configuration Control Sheet

Date	Filename/version #	Author	Revision Description
8/24/2005	*Releases and Capabilities, version 1.0*	FDsys team	First Draft for P&S review
8/30/2005	*Releases and Capabilities, version 1.1*	Gil Baldwin	Version with matrix corrections
9/9/2005	*Releases and Capabilities, version 2.0*	Mike Wash	Added Change / Configuration Chart
9/28/2005	*Releases and Capabilities, version 2.0*	Lisa LaPlant	Changed Version Control per Comments from the Team Review of the Draft Version Control Specification.

Contents

FDsys System Releases and Capabilities

Executive Overview

GPO is responsible for the preservation and dissemination of information products generated by the entire Federal government, including current, legacy and future products in all conceivable formats.

GPO's Future Digital System (FDsys) will provide a comprehensive, systematic and dynamic means for preserving electronic content free from dependence on specific hardware and/or software. The system should automate many of the electronic content lifecycle processes and make it easier to deliver electronic content in formats suited to customers' needs.

Standing up the FDsys is a complex system integration task, which will be rolled out in a series of releases. Each release includes improvements to both system capability and underlying infrastructure, and is built incrementally on those preceding it, until the full range of capabilities is implemented.

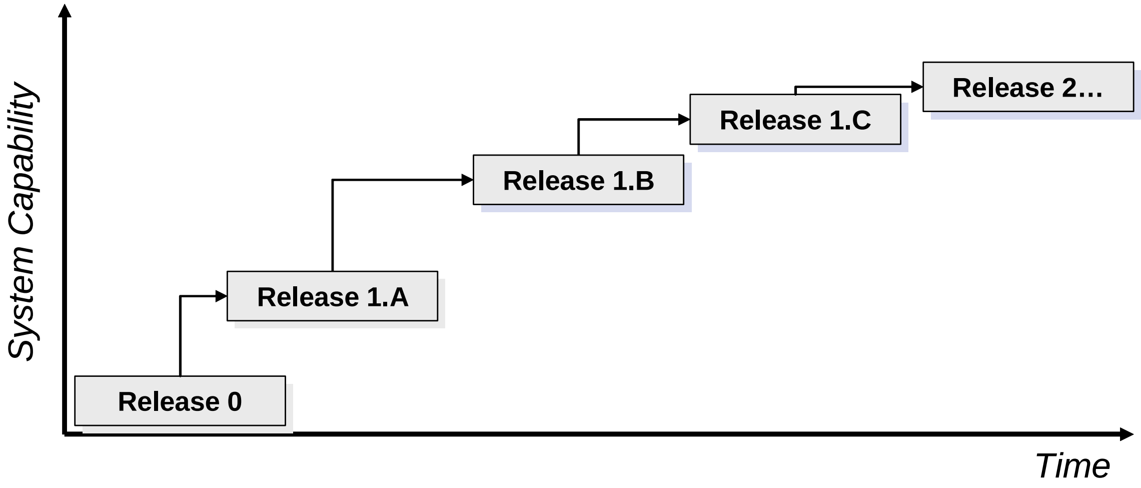
In this document, the FDsys releases are presented at a high level. The planned release schedule begins with Release 0 (zero) in late 2005, intended to support content submitted by Digital Conversion Services; Releases 1.A, 1.B and 1.C to deliver the basic functionality by late 2007, and Releases 2 and 3 to deliver enhancements continue through 2008

The document breaks down the functionality and infrastructure descriptions into seven clusters. Each cluster is described and then detailed by release.

Additional supporting detail is available in:

- *GPO Concept of Operations for the Future Digital System (ConOps v2.0),* Final, May 16, 2005.
- *Requirements Document for the Future Digital System (RD v1.0),* Final, May 16, 2005.

FDsys Schedule of Releases



Release 0 - Target Date October 2005
Supporting Digital Conversion Services for GPO Access

Release 1.A - Target Date July 2006
Content Submission and Basic Infrastructure

Release 1.B – Target Date January 2007
Access and Content Delivery Enabled

Release 1.C - Target Date July 2007
Basic Functionality Completed

Release 2 - Target Date January 2008
Enhanced Capabilities

Release 3 – Target Date July 2008
Enhanced Access Tools

FDsys Clusters Overview

Features and Functionality

I. Style Tools

II. Content Submission
   A. Deposited Content
   B. Harvested Content
   C. Converted Content
   D. Content Originator Ordering

III. Content Management
   A. Unique ID
   B. Persistent Name (Handles)
   C. Authentication
   D. Version Control
   E. Data Mining

IV. Content Access
   A. Search
   B. Cataloging/Reference Tools
   C. Request
   D. Interface
   E. User Support

V. Content Delivery
   A. Hard Copy
   B. Digital Media
   C. Electronic Presentation

VI. Content Preservation

System Infrastructure

VII. System Infrastructure
   A. Workflow
   B. Security
   C. Storage
   D. Enterprise Service Bus
   E. Content Management Suite

Capabilities, Functions, & Features by Release (through Release 2)

		Release 0	Release 1A	Release 1B	Release 1C	Release 2
		October 2005	July 2006	January 2007	July 2007	January 2008
		Supporting Digital Content Services for GPO Access	**Content Submission & Basic Infrastructure**	**Access & Content Delivery Enabled**	**Complete Basic Functions**	**Enhanced Capabilities**
I. Style tools			Initial functionality (Microcomp replacement)	Additional functionality	Additional functionality	Enhanced collaboration tools
II. Content Subm.	**A. Deposited Content**		Ingest capability, package creation			Custom composition
	B. Harvested Content		Ingest capability, package creation	Harvester	Enhanced harvester	Collect integrity marks
	C. Converted Content	Accept CCPs - create GPO Access Packages	Ingest capability, package creation			Process encrypted files
	D. Content Originator Ordering	(Existing legacy systems)			Basic Functionality	Enhanced capabilities
III. Content Mgt.	**A. Unique Identifier**	Assign for CCP & jobs	Assign & manage for all content		Unique ID for content ordering	
	B. Persistent Name	Assign to GPO Access Packages	Assign & manage for all content			
	C. Authentication	Authenticate PDFs	Validate & manage			
	D. Version Control	Accept & transfer version information	Auto detection / trigger	Determine and record relationships between versions		Enable rejection of duplicate content
	E. Data Mining		Support security and admin functions		Basic functionality	Customized reports

Capabilities, Functions, & Features by Release (through Release 2) – cont.

		Release 0	**Release 1A**	**Release 1B**	**Release 1C**	**Release 2**
		October 2005	July 2006	January 2007	July 2007	January 2008
		Supporting Digital Content Services for GPO Access	**Content Submission & Basic Infrastructure**	**Access & Content Delivery Enabled**	**Complete Basic Functions**	**Enhanced Capabilities**
IV. Content Access	**A. Search**	GPO Access search functionality		Basic functionality	Enhanced search features	Customizable access tools
	B. Cataloging & Reference Tools	Full cataloging (ILS) functionality		Generate reference tools from metadata	Customize reference tools	
	C. Request	GPO Access functionality		Basic functionality	End user ordering	
	D. Interface	GPO Access functionality	UI for all functions of this release	UI for all functions of this release	UI for all functions of this release	UI with customization
	E. User Support	ID contact center	US for all functions of this release	US for all functions of this release	US for all functions of this release	Interactive US
V. Content Delivery	**A. Hard Copy**			Basic functionality	Enhanced functionality	
	B. Digital Media			Basic functionality	Enhanced functionality	
	C. Electronic Presentation	GPO Access Package / Sec. 508 compliant		Basic functionality	Enhanced functionality	
VI. Content Preserv.			Preservation storage (repository environment)		Preservation processes	

Capabilities, Functions, & Features by Release (through Release 2) – cont.

		Release 0	**Release 1A**	**Release 1B**	**Release 1C**	**Release 2**
		October 2005	July 2006	January 2007	July 2007	January 2008
		Supporting Digital Content Services for GPO Access	**Content Submission & Basic Infrastructure**	**Access & Content Delivery Enabled**	**Complete Basic Functions**	**Enhanced Capabilities**
VIII. Sys Infrastructure	**A. Workflow**		Workflow to support functionality in this release	Workflow to support functionality in this release	Workflow to support functionality in this release	TBD
	B. Security		Security to support functionality in this release; infrastructure security	Security for delivered content	Ordering, item into financial and inventory system (Oracle)	TBD
	C. Storage	Store CCPs; WIP storage for GPO Access Packages	Supporting WIP, ingest, ACP, AIP	DIP storage; records management for temporary stores	Data mining	TBD
	D. Enterprise Service Bus		Application integration; integration of ILS	Continued application integration	Integrate Oracle for financial and inventory management	TBD
	E. Content Management Suites	Initial functionality for CCPs	Support for ingest, content submission, content management, content access functions, and preservation	Support for harvester, additional style tool capability, version control, content delivery, authentication, records management	Support for enhanced harvest, CO order, access, style tools, user support, and support for data mining, and ordering	TBD

Cluster 1: Style Tools

The FDsys must provide a simple, easy-to-use method to encourage Content Originators to provide content to GPO. This will be accomplished by the design of a collection of tools that allows for the seamless creation, submission, validation, management, and subsequent access to the content.

Therefore, Style Tools will be a tool or suite of tools that allows content creators to easily create and manage their content so that GPO can readily ingest the content into the FDsys.

The purpose of a style tool is to allow GPO to move further upstream in the publishing process - closer to content creators and the content the FDsys needs. This will allow GPO to capture more content, thus minimizing fugitive documents.

	Release 0	Release 1.A	Release 1.B	Release 1.C	Release 2
I. Style tools		●	●	●	●

Release 0

- None in this release

Release 1.A

- Provide both the ability to accept content from alternate sources and create content in stand alone and collaborative processes.
- Incoming digital objects must be assigned unique IDs.
- Pre-defined templates or other aesthetic driven tools will compose content as desired by the content originator.
- The content will be presented visually for inspection (e.g., proofs) through hard and soft copy displays with approval rights granted as appropriate.
- Once the content is finalized and approved, Style tools will pass a valid content package to the system.

Releases 1.B – 1.C

Subsequent releases will build on and enhance the core capabilities from release 1.A. These include adding enhanced graphic editing and creation capabilities, artificial intelligence based composition tools and other customer based tools for making content creation more efficient.

Cluster 2: Content Submission

Digital content in the Future Digital System will come from three sources: legacy content converted to digital form (digitized), digital content harvested from the World Wide Web, and born digital content pushed to GPO (deposited) by its originator. While the last of these, deposited content, will eventually be the predominant source, the harvested and converted content will be significant in the early life of the system, and will continue to play a role into the near future.

Deposited content is defined as content intentionally submitted to GPO by Content Originators. The SIP for deposited content will include the digital object received from the Content Originator as well as corresponding customer processing requirements and additional metadata.

Harvested content is content within the scope of dissemination programs that is gathered from Federal agency Web sites. Discovery, assessment, and harvesting tools will be used to create a SIP.

Converted content is digital content created from a tangible product. GPO will continue to work with various user communities on digitizing a comprehensive collection of legacy materials. This digital collection will be made available in the public domain for permanent public access through GPO's dissemination programs.

Content Originator (CO) Ordering provides a system interface for Content Originators. Content Originators may submit content, order and re-order content, and specify Content Delivery and other service options through CO Ordering. CO Ordering should allow Content Originators and GPO Users to discover the cost of content delivery, choose delivery options, request delivery, and discover payment/billing status for delivery of content when applicable.

		Release 0	Release 1.A	Release 1.B	Release 1.C	Release 2
II. Content Submission	A. Deposited Content		•			•
	B. Harvested Content		•	•	•	•
	C. Converted Content	•	•			•
	D. Content Originator Ordering				•	•

Deposited Content

Release 0

- None in this release

Release 1.A

- All deposited content will be accepted by the system
- All deposited content will be assigned a unique identifier
- Deposited content will be checked for integrity and completeness
- Excepted files (those with encryption/usage restrictions) will be identified by the system
- Deposited content files will have basic descriptive, structural, and administrative metadata upon submission,
- Deposited files and metadata will be stored upon submission

Release 1.B

- None in this release

Release 1.C

- Deposited Content to be sent through style tools before ingest, allowing use of a template to assist content providers in pulling together information for a publication.

Release 2

- Encrypted files may be processed through an alternate workflow
- Deposited Content originators will be notified when content has been received for submission

Release 3

- None in this release

Harvested Content

Release 0

- None in this release

Release 1.A

- The ability to accept all content and metadata delivered by the harvesting function into WIP (Work-in-Process) Storage

• The ability to create a Submission Information Package from the harvested content and metadata

• The ability to discern whether harvested content or metadata needs to be ingested (e.g., is it already in the system?)

Release 1.B

• Capability to discover, assess, characterize, and collect content and available metadata from Federal agency Web sites that fall within the scope of GPO dissemination programs.

• Capability to determine if the discovered content is within the scope of GPO dissemination programs.

• The ability to locate and collect all file types that may reside on Content Originator Web sites

• Capability to collect content in the exact form that the content was resident on the agency Web site.

• Capability to accept and apply rules and instructions that will be used to assess whether discovered content is within the scope of GPO dissemination programs

• Capability to provide quality control functions to test accuracy/precision of rule application and to incorporate results into rule creation/refinement.

• Capability to produce reports on harvesting activities

Release 1.C

• The ability to harvest deep Web information within, on, or behind query-based databases, agency content management systems, dynamically generated Web pages, FTP servers, proxy servers, firewalls.

Release 2

• The harvester shall have the ability to collect integrity marks associated with content as it is being harvested.

Release 3

- None in this release

Converted Content

Release 0

- The system must accept Converted Content Packages (CCPs) from all sources.
- The system must create GPO Access Packages following the *GPO Access* specification for direct accessibility on *GPO Access*.

Release 1.A

- The capability to accept digital content created by conversion processes (e.g., scanning, text encoding).
- The capability to check provided digital file(s) (e.g., Virus Check, checksum) prior to ingest
- The capability to accept version information and integrity marks from converted content
- The capability to accept files which contain all agency processing information with the converted content, including billing information, jacket number, agency reference information, etc.
- The to record basic descriptive, structural, and administrative metadata
- The ability to convert tangible titles that have rights limitations, if they are in scope for GPO dissemination programs
- The ability to ingest content packages from Release 0 stores for use in subsequent releases of the Future Digital System

Release 1.B-C

- None in this release

Release 2

- The ability to process encrypted files through an alternate workflow to obtain key information to allow the file to be opened.
- The ability to provide notification to the submission agency/authority that the content has been received.

Release 3

- None in this release

Content Originator Ordering

Major CO Ordering functionality will:

- Provide Content Originators with a direct interface to the system.
- Allow Content Originators to order and reorder content.
- Capture all agency processing requirements (e.g., the system must capture all relevant metadata the Content Originator supplies).
- Provide the cost of content delivery based on Content Originator order information.
- Allow GPO Service Specialists to augment Content Originator orders and job specifications (e.g., riders).
- Allow Content Originators to specify options (e.g., output media, quantities, specifications, other preferences) for delivery of content (hard copy, electronic presentation, and digital media).

Release 0

- CO orders for converted content are handled using existing systems.

Release 1.A-B

- None in this release

Release 1.C

- Most CO Ordering functions roll out in this release.

Release 2

- The system shall provide the capability for Content Originators to select from GPO's approved external Service Providers.
- The system shall provide response back to an order request within a timeframe established by GPO business units responsive to C.O. need.
- The system shall provide C.O. specific static and dynamic summary reports for various components of the system including Service Provider performance, Content Originator activity, response times, product types, dollar values/totals, number of jobs awarded to individual Service Providers.

Release 3

- None in this release

Cluster 3: Content Management

The Content Management cluster includes five functions: Unique Identifiers, Persistent Name, Authentication, Version Control, and Data Mining

Unique Identifiers are character strings that uniquely identify all content within FDsys throughout the content lifecycle. The Unique ID will be assigned once, and will be indefinitely linked to the corresponding content.

Persistent Name provides for the identification of each digital object independent of its location. It will provide permanence of identification, resolution of location, and globally registered, validated, and unique.

Authentication will assure customers that the information made available by GPO through the Future Digital System is authentic and/or official. There is a need for information that is reliable because it is from a trusted source, and a need to ensure the protection of data against unauthorized modification or substitution of information.

Version Control will evaluate and establish the version of a piece of content and subsequently track it through its entire life cycle, based on best practices. Version control will be called upon to analyze Content Packages and assign the appropriate version, consistent with requirements for version triggers and chain of responsibility. The chain of responsibility will be reflected in the metadata.

Data Mining is associated with finding, aggregating, and associating business process information through the analysis of collections of data.

		Release 0	Release 1A	Release 1B	Release 1C	Release 2
III. Content Management	A. Unique Identifier	•	•		•	
	B. Persistent Name		•			
	C. Authentication	•	•			
	D. Version Control	•	•	•		
	E. Data Mining		•		•	•

Unique Identifiers

Release 0

- Assign unique ID for converted content

Release 1.A

- Unique ID functions will assign an alphanumeric identifier for each unique digital object.
- Unique ID's are truly "unique" and will only be used once per digital object.
- Unique ID information will be recorded in metadata.

Release 1.B

- None in this release

Release 1.C

- Unique ID functions will assign an alphanumeric identifier for each unique job.
- Unique ID for unique jobs will be used only once and recorded in metatadata.

Persistent Name

Release 0

- None in this release

Release 1.A

- All versions of content packages will be assigned persistent names when ingested into the system
- Persistent Names will be human readable and interoperable with legacy schemes and other systems
- Persistent Names will be location independent, and a resolution system will locate and provide access to the content.
- Persistent Name information will be recorded in metadata, and all Persistent Name transactions will be logged.

Authentication

Release 0

- Authenticate PDFs derived from converted content.

Release 1.A

- Authentication will certify content as authentic and/or official at both document and granular levels.
- Authentication will verify and validate content that is deposited, harvested and converted at ingest.
- Authentication will notify when changes were made to content, where the changes were made to content, by who the changes were made to content, and what changes were made to content.
- Authentication will use integrity marks that convey authentication information to users.

Version Control

Release 0

- Version identifiers for CCPs will be detected when they are present, and assigned when they are not present.

Release 1.A

- Version Control will provide for a record of the chain of responsibility of content and the system will manage multiple versions of works in progress.
- Version Control will express version information in metadata.
- Version Control will enable rules to be applied for version triggers.
- Version Control will support version detection for submission information packages.
- Version Control will support the detection of version triggers, such as modifications to content, changes in file format, changes in a publication’s title.

Release 1.B

- The system shall determine and record relationships between versions.

- Version Control will apply rules for version triggers to groups of related content.
- Version Control will support notification to GPO when version triggers are activated and when version information cannot be determined.

Release 1.C

- None in this release.

Release 2

- None in this release.

Release 3

- Version Control will support the detection of changes that were previously detected and excluded as a version trigger.

Data Mining

Release 0

- None in this release

Release 1.A

- Data Mining will support security and administration functions, such as the logging of transactions and the creation of metadata.
- Data Mining will support the storage of extracted data.

Release 1.B

- None in this release

Release 1.C

- Data mining will support the extraction, analysis, and presentation of business process information.
- Data mining will provide for data extraction by format, random samples of data, and other parameters.
- Data mining will provide for data presentation and data interface through reports showing for instance, trend analysis and usage patterns and the export of results.
- Data mining will support data analysis and data modeling.

Release 2

- Data mining will support the creation of customized reports for analysis.
- Data mining will support more customization of views of data based on user preferences.
- Data mining will provide notification alerts to users based on defined criteria.

Release 3

- Data Mining will incorporate taxonomies and ontologies.

Cluster 4: Content Access

Access will be the primary interface between End Users and the FDsys. Access has been divided into five functional areas:

- Search – Performing queries on content and metadata so that content may be retrieved from storage and delivered to users.
- Cataloging and Reference Tools – Adding metadata to content in the form of standard bibliographic records and lists and resources that assist users in locating and accessing content.
- Request – Requesting delivery of content and metadata.
- Interface – Creating user and system interfaces for all functional areas, as needed.
- User Support – Supporting access to content and services.

Summary of Cluster 4 Capabilities by Release

0 **(target 10/05)**	**1.A** **(target 7/06)**	**1.B** **(target 1/07)**	**1.C** **(target 7/07)**	**2** **(target 1/08)**	**3** **(target 7/08)**
Basic Cataloging via the ILS	Core Access capabilities Enhanced Cataloging including Metalib and SFX Default Interfaces Basic User Support for enabled features including Cataloging	Basic Reference Tools Basic Search Basic Request Default Interfaces Basic User Support for enabled features including Search, Reference Tools, and Request	Enhanced Search Integration with legacy and external repositories Enhanced request with e-commerce features enabled and integration with financial and inventory systems Enhanced User Support Customized Reference Tools	Basic user customized Search Automatic creation of metadata and dynamically generated Reference Tools Enhanced Interfaces with basic user customization Interactive training and User Support Customized Request	Fully customized and personalized Search, Request, Interface, Reference Tools, User Support and training that is based on user preferences Integrate Search, Reference Tools, Request, and User Support seamlessly into a single interface

Core Access Capabilities

Release 1.A

- Provide open and interoperable access to Content Packages in the system.
- Provide access to in-scope content not resident within the system.
- Create and provide access to section 508 compliant Access Content Packages.
- Allow GPO users to designate levels of access based on user privileges and credentials.
- Perform records management functions for Access.

Search

Release 0 - 1.A None

Release 1.B

- Provide basic search capabilities.
- Search and retrieve all available metadata and content collections (internal and external) both simultaneously and separately based on access rights, privileges, and GPO business rules.
- Provide basic search functionality such as Boolean search language, fuzzy logic, natural language, proximity, and synonyms searching.
- Apply multiple search logic in a single search query (e.g., Boolean, truncation, wildcards, nesting).
- Perform a search for conceptually related terms (e.g., search for "World Series" returns articles on the Red Sox).
- Provide for different search complexity levels (e.g., simple search, advanced/fielded search) and multiple search interfaces based on search skill level and user class.
- Provide, at a minimum, search results and the ability to sort results lists by title, date, and relevancy.
- Display, at a minimum, title, file size, version, content collection (source), and an identifier (link).
- Provide the ability for users to refine search queries.
- Log search transactions (e.g., user access, administrative changes).

Release 1.C

- Provide enhanced search features.
- Deliver search results at the finest level of granularity supported by the target content package.
- Sort results by content collection, format, version, and defined metadata fields.
- Perform federated searches across multiple internal and external repositories, including legacy repositories (e.g., WAIS).

Release 2

- Provide basic user customized search features.
- Provide the ability for users to establish a user profile to store searches, preferences, and results sets.
- Automatically execute saved searches and deliver search results to users.

Release 3

- Provide fully customized and personalized search options, interfaces, and results lists that are based on user preferences.
- Analyze search results.
- Cluster search results.
- Display search results graphically.

Cataloging & Reference Tools

Release 0

- Virtually all of the capabilities listed for the cataloging function are a part of the ILS and are included in FDsys.
- GPO has acquired Oracle-based, COTS, cataloging software, the Aleph 500 system from Ex Libris, USA, also referred to as the Integrated Library System (ILS). The FDsys must interface with the ILS and the Online Computer Library Center, Inc. (OCLC) system for creating bibliographic metadata. Bibliographic metadata includes links to content maintained on various sites managed by or completely external to GPO, e.g. the OCLC Digital Archive, on various agency and library sites, etc.

Release 1.A

- None in this release

Release 1.B

- Provide enhanced Cataloging features.
- Generate reference tools (lists) based on any indexed metadata field, and to link to external content and metadata. These lists may be static pages produced from report generation capabilities, or dynamic results lists from searches. These searches may be pre-configured ("canned") or individually created for one-time use.
- Create, acquire, and store metadata (e.g., MARC), references to metadata (e.g. Subject Bibliographies), and references to content (e.g., Federal Agency Internet Sites, Browse Topics, etc.). Reference Tools will include lists and resources that assist users in locating and accessing content.

Release 1.C

- Generate and manage customized reference tools including lists of content available for sale, for selection by depository libraries, content received by a specific library, related resources, etc.

Release 2

- Automatically create metadata.
- Link related resources in descriptive metadata.
- Provide interoperability with third party reference tools (e.g., search catalogs of other libraries).
- Dynamically generate Reference Tools.

Release 3

- Provide fully customized and personalized Reference Tools that are based on user preferences.
- Generate lists based on user preferences, and support interactive processes so users can create Reference Tools.

- Deliver customized Reference Tools in both electronic and tangible formats.

Request

Release 0 - 1A

- None in this release

Release 1.B

- Provide basic request and delivery options for internal and external content and metadata as defined by GPO business units.
- Select format and file types of requested content from available options.
- Select level of granularity for requested content from available options.

Release 1.C

- Provide enhanced request with e-commerce features enabled.
- Integrate with financial and inventory systems.
- Generate pricing information for the delivery of content.
- Accept payment for content.
- Enable shopping cart functionalities.
- Create unique order numbers for each request.
- Allow secure and encrypted ordering.
- Allow third party ordering.
- Generate confirmations and transaction receipts.
- Accommodate rider orders from GPO, Federal Agencies, and Congress.

Release 2

- Provide customized request and delivery options.
- Support custom composition and content formatting from available options.
- Store pending orders in a shopping cart feature.
- Store and access user preferences and request histories in a secure environment (e.g. request status, delivery preferences, preferred payment methods, request tracking, prior request history).
- Provide lists of new and popular titles, best sellers, and other lists as defined by GPO business rules.

Release 3

- Provide fully customized and personalized request options are based on user preferences.
- Deliver personalized offers based on user request history (e.g. "you may also be interested in….")

Interface

Release 0

- Use existing GPO Access interface

Release 1.A

- Provide Interfaces for Release 1.A functions.
- Provide a default workbench (set of available tools) for each user class.
- Provide a default workbench for public end users that does not require them to log-in.
- Display the appropriate default workbench based on a user's rights and privileges.
- Allow users or groups of users to login and create an account.
- Provide user interfaces capable of rendering supported types of electronic files.
- Provide system interfaces that promote interoperability among networked systems (e.g., APIs).
- Comply with best practices and guidelines regarding usability for interface (e.g., The Research-Based Web Design & Usability Guidelines function as best practices for Federal Agencies).
- Comply with accessibility laws and regulations (e.g., Section 508 of the Rehabilitation Act Amendments of 1998) as appropriate.
- Conform to current World Wide Web Consortium (W3C) guidelines for interoperable technologies.

Release 1.B

- Provide Interfaces for Release 1.B functions.

Release 1.C

- Provide Interfaces for Release 1-C functions.

Release 2

- Enhanced interfaces with basic user customization including:
  - Selecting available tools.
  - Customizing the appearance of an interface.
  - Creating interfaces for subsets of user classes.

Release 3

- Provide fully customized and personalized interfaces based on user preferences.
- Integrate search, reference tools, request, and user support seamlessly into a single customizable interface.

User Support

Release 1.A

- Provide basic support for enabled features including Cataloging.
- Provide multiple methods of contact for user assistance (e.g., web, phone, email, U.S. mail, chat)
- Allow users to query a knowledge base.
- Allow users to submit questions to a helpdesk.

- Automate routing of inquiries to the departments/individuals according to workflow guidelines
- Support priority processing (e.g., alerts, assignments to staff, nonstandard workflow, tracking, etc.)

Release 1.B

- Provide basic support for enabled features including search, reference tools, and request.
- Provide context specific help on all user interfaces (e.g. help functions related to what is being viewed.)
- Provide alert services which automatically deliver information about content based, and allow users to subscribe and unsubscribe to the alert services.
- Provide users with access to their questions and responses and related questions and responses.
- Provide automatic confirmation of receipt of questions/inquiries.
- Create reports on user support activities.

Release 1.C

- Provide enhanced support.
- Provide online tutorials.
- Provide customized services for user classes and sub-groups within user classes as defined by GPO.
- Manage an unlimited number of user records and queries.

Release 2

- Support real-time, interactive information exchange (e.g., chat, discussion groups, web conferencing).
- Provide interactive training applications.
- Allow users to enroll in training and other events and access training materials and training history.
- Allow GPO users to manage training and events.
- Perform records management functions on knowledge base data, user information, and exchange logs.

Release 3

- Provide fully customized and personalized support and training that is based on user preferences.
- Allow users to measure their progress and performance within online training applications.
- Allow users to provide feedback about online training.

Cluster 5: Content Delivery

Content Delivery refers to mechanisms for delivering content in a method or manner that fits the requirements of the Content Originator or End User. The delivery methods identified are hard copy, electronic presentation, and digital media.   Content Delivery requirements reflect the expansion of electronic presentation as the primary dissemination method for authentic and/or official content of the Government.

		Release 0	Release 1.A	Release 1.B	Release 1.C	Release 2
V. Content Delivery	A. Hard Copy			•	•	
	B. Digital Media			•	•	
	C. Electronic Presentation			•	•	

Release 0 and 1.A

- None in these releases. Delivery of electronic content continues via GPO Access.

Release 1.B

• The system will retrieve Access Content Packages (ACPs) from storage based on an End User request. The ACP is transformed into a Dissemination Information Package (DIP) for delivery to the end user. DIPs may include integrity marks certifying the content as official and/or authentic. DIPs are created for:

  - o Transfer to a Service Provider for Hard Copy output,
  - o Delivery to an End User in electronic format,
  - o Delivery to an End User via Digital Media (e.g., CDs, DVDs), and
  - o Delivery to Digital Media (devices external to the system such as PDAs, MP3 players, etc.).

- The system will be able to generate DIPs for the most cost effective method for Hard Copy output (e.g., offset press, digital printing). The system will support Print on Demand.

- The system will generate Electronic Presentation DIPs for multiple computer platforms (e.g., Windows, Macintosh, UNIX) and non-desktop electronic devices (e.g., PDAs, MP3 players, e-books). Multiple formats will be available for delivery to End Users.

- The system will generate DIPs for transfer via digital media (e.g., CDs, DVDs) and transfer to digital media (devices such as PDAs, etc.)

- Content may be pushed from the system to users (e.g., RSS feeds, e-mail, raw data feeds) and pulled from the system by users (e.g., FTP).

- Notification of content delivery can be provided electronically to the End User and to GPO. Options can be provided to End Users when content is not delivered (e.g., resubmit content, cancel delivery, etc.)

- A database of Service Providers will be managed by GPO. This database will allow Service Providers to electronically maintain their profiles (e.g., equipment, capabilities) and will allow authorized GPO users to manage Service Provider performance information (e.g., quality and performance data). Content Originators and users will be able to select Service Providers based on their needs. Tracking of orders and logging of Service Provider activities (e.g., job received, job complete) will also be possible.

- The system will comply with GPO Quality Assurance standards for content delivery, best practices and guidelines regarding usability for electronic content (e.g., The Research-Based Web Design & Usability Guidelines function as best practices for Federal Agencies), accessibility laws and regulations (e.g., Section 508 of the Rehabilitation Act Amendments of 1998), and W3C guidelines.

Release 1.C

- None in this release

Release 2

- The system shall have the capability to deliver DIPs to non-GPO storage devices (e.g., customer servers).

Release 3

- None in this release

Cluster 6: Content Preservation

Preservation enables comprehensive, timely, permanent public access to the final published, official version(s) of U.S. Government publications in digital formats, by the retention of faithful, fully functional master files of content, and the performance of processes which assure ongoing usability of those files.

	Release 0	Release 1.A	Release 1.B	Release 1.C	Release 2
VI. Content Preservation		•		•	

Release 0

- None in this release

Release 1.A

- Preservation copies of digital publications, called Archival Information Packages (AIPs), with associated technical metadata, will be maintained in FDsys Archival Storage.
- Release 1.A establishes the basic system infrastructure to keep content alive, by preserving AIPs. In order of preference, the outcomes desired are:
    - Fully, faithfully duplicated files, rendered using the original application.
    - Files which faithfully reproduce content, behavior and appearance of the original, rendered using other software than the original application.
    - Files which exactly convey the content but may alter behavior and/or appearance, rendered using other software than the original application.
- The system shall preserve all essential behaviors of digital content, including content functionality and presentation, context, and structure.
- The FDsys digital archival repository environment will be based on open-standards architecture.

Release 1.B

- None in this release

Release 1.C

- Release 1.C includes the preservation processes the FDsys will use to assess the condition of digital content and initiative the necessary actions to keep that content usable. The preservation process employed in any given situation should be the least intrusive; i.e. that which alters the original AIP the least.
- The specific preservation processes required by GPO are a policy determination. FDsys must be capable of supporting activities necessary to keep content accessible and usable, including
    - Migration
    - Refreshment
    - Emulation
- Management of preservation processes include:
    - File backup/redundant storage
    - Establishing the duration of preservation
    - Creating replacement content packages
- File disposition options include:
    - Transfer to NARA
    - Scheduled retirement of selected content

Release 2-3

- None in these releases

Cluster 7: System Infrastructure

The System Infrastructure Cluster includes the system elements of Workflow, Security, Storage, Enterprise Service Bus (ESB) and Content Management Suites (CMS).

Workflow is the automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules. A workflow management system defines, creates and manages the execution of workflows through the use of software, running on one or more workflow engines, which is able to interpret the process definition, interact with workflow participants and, where required, invoke the use of IT tools and applications

Security includes the protection of systems (applications) against unauthorized access to or modification of information, whether in storage, processing or transit, and against the denial of service to authorized users or the provision of service to unauthorized users, including those measures necessary to detect, document, and counter threats. The measures and controls, including physical controls in conjunction with management, technical and procedural controls, that ensure the confidentiality, integrity and availability of information processed and stored by a system.

Storage includes the functions associated with saving digital publications on physical media, including magnetic, optical, or other alternative technologies. Akamai is currently being used.

ESB is a new approach to integration that can provide the underpinnings for a loosely coupled, highly distributed integration network that can scale beyond the limits of a hub-and-spoke enterprise application integration broker. ESB is a standards-based integration platform that combines messaging, web services, data transformation, and intelligent routing to reliably connect and coordinate the interaction of significant number of diverse applications across extended enterprises with transactional integrity.

Content Management Suites manage information throughout its life cycle, from creation to archiving. Enterprise Content Management and Content Management suites support regulatory compliance, risk management, customer self-service, CRM applications, and vertical processes. These suites support such system functionality as document services, records management, forms, authoring, imaging, transformation, import/export, indexing, storage management, file systems, workflow, library services, search, collaboration, development kits, publishing, source storage, integration, etc.

		Release 0	**Release 1A**	**Release 1B**	**Release 1C**	**Release 2**
VII. System Infrastructure	A. Workflow		●	●	●	●
	B. Security		●	●	●	●
	C. Storage	●	●	●	●	●
	D. Enterprise Service Bus		●	●	●	●
	E. Content Management Suites	●	●	●	●	●

Workflow

Release 1.A
Workflow software that supports enabled features such as:

- Ingest (converted, harvested and deposited content)
- Unique ID, Persistent Name, Authentication, Validation, Version Control, Cataloging, AIP, ACP generation, preservation storage, initial style tools (MCR), UI and User support.
- Automation of workflow processes that were previously manual.

Release 1.B

- The Harvester enabled, additional style tool capability, auto version trigger and detection, search request, content delivery, UI and user support for enabled features

Release 1.C

- Support of additional features--enhanced harvester, CO ordering, Unique ID for Job Orders, customizable access tools, reference tools, enhanced user support, data mining, enhanced style tools, request to support transactions (ordering), preservation processes

Release 2-3

- TBD

Security

Release 1.A
Oriented to Content Submission and FDSys system functions. It is assumed that the user base consists of GPO administrator personnel and content submitters.

- Access Control to FDSys objects by user ID
- Authentication
    - User authentication – each user unique identifier
- Audit Logs
    - Logs of system activities
- Security Administration
    - Administration tools and capabilities present to allow efficient administration of security capabilities
- Availability
- Integrity
    - Methods to ensure data is not accidentally or intentionally altered
- Confidentiality
    - Data is not readable by unauthorized parties as appropriate (example, user passwords)
- Privacy
- Standards Based

Release 1.B
Oriented to capabilities to support functions introduced for Content Delivery
<u>Assumption:</u> User base expands to include general users and content acquisition by 3rd parties.

Release 1.C

- Audit Logs for Data Mining
- Access Control limitations to data mine appropriate data
- Temporary storage and security to support data mining

Release 2-3

- TBD

Storage

Release 1.A

- Storage system to support content ingest
- WIP storage (high availability/high access)

- ACP storage (moderate availability/access, backup)
- AIP/Preservation storage (low access/availability, backup)
- Storage management.

Release 1.B

- Resizing of required storage
- ACP storage (high availability/access, backup)
- Temporary DIP
- Records management for temporary stores

Release 1.C

- Temporary storage and security to support data mining

Release 2-3

- TBD

Enterprise Service Bus (ESB)

Release 1.A

- Integration of ESB platform to enable application integration
- Integration of the ILS

Release 1.B

- Continued application integration

Release 1.C

- Tie into Oracle for financial and inventory management

Release 2-3

- TBD

Content Management Suites

Release 0

- Functionality to support the standup of Digital Conversion Services' production and management of Converted Content Packages.

Release 1.A
Content management solutions to support:

- Ingest (converted, harvested and deposited)

- Unique ID, Persistent Name, Authentication, Validation Version Control, Cataloging, AIP, ACP generation, Preservation storage, Initial Style Tools (MCR), UI and User Support for enabled features

Release 1.B

- Content management solutions that support the enabled Harvester, additional style tool capability, auto version trigger and detection, search request, content delivery, UI and user support for enabled features
- Authentication for delivered content
- Records management for temporary stores

Release 1.C

- Support of additional features--enhanced harvester, CO ordering, Unique ID for Job Orders, customizable access tools, reference tools, enhanced user support, data mining, enhanced style tools, request to support transactions (ordering), preservation processes

Release 2-3

- TBD

System Specifications and Characteristics

(October 3, 2005)

Specifications and Metrics for the Future Digital System (FDsys)

Version 1.0
Last saved: 10/3/2005 3:34 PM

United States Government Printing Office (GPO)

Document Change/Configuration Control Sheet

Date	Filename/version #	Author	Revision Description
10/3/2005	*System_Sizing*	FDsys team	V 1.0

1 Scope

1.1 Overview

Each of the functional elements of the system will have independent specifications to support development. However, there are core system characteristics that either do not support a specific functional element or which cross multiple or all system functions. These core characteristics are outlined in this document.

The core system characteristics have been divided into the following categories:

Order Stats
Content Size and Characteristics
Web Traffic Statistics for Delivery
Content Delivery Statistics
Data Location
User Support
    Call Center
    Customer Services
Number of Users

2 Referenced Documents

2.1 GPO

ConOps 2.0
RD 1.0
Statistics Book for FY2004 (GPO Internal Document)

2.2 Government

2.3 Organizational/Standard

3 Current Situation

There is no current single system configuration; therefore, current statistics are an amalgamation of individual reports. These reports are generated from Mainframe applications, client/server applications and standalone desktop applications. Internal network traffic is deemed reliable; however, external traffic is difficult to gauge since GPO does not employ cookies, but relies on unique IP addresses.

Content Delivery statistics are based on current tangible titles delivered, and do not reflect electronic delivery.

Call Center and Customer service “Incident creations” (phone and email) are segregated by business unit. Therefore the data may not accurately reflect the current state and are provided as best estimates.

Number of users by user class data should be reliable in this document.

3.1 Current Situation and Metrics

Order Stats: are handled independently as “black” or “red” jackets (job orders). Red is a procured order and Black are in-house production orders. Jacket numbers (red and black) are 3 numeric characters separated by a “-“ and followed by 3 numeric characters (e.g., xxx-xxx). These data reflect only the unique and initial orders submitted to GPO by the ordering agency.

- In-Plant – 16,410 orders in FY2004
- Procured – 123,229 orders in FY2004
- Total – 139,639 unique orders placed by agency customers in FY2004

NOTE: Projections for unique orders is Flat based on Customer Services budget estimates for FY2006.

Content Size and Characteristics of SIPs:
Deposited Content: Average size (in MB) of files submitted to GPO is 120 MB[1] in FY2004
Harvested Content: Average size (in MB) of files submitted to GPO is 10 MB[2] in FY2004
Converted Content: Average size (in MB) of files converted by GPO is 360 MB[3] in FY2004

Web Traffic Statistics for Delivery of Electronic Content
Document Retrievals per month:

- 14 million per month from WAIS database
- 6 million per month served through Akamai
- 6 million per month served through Web servers
- 10 million per month served through Web servers on Hosted Sites

Visits Transactions and Retrievals: 417,160,156 per year[4]

Pages Produced Through GPO Services
Pages (procured) [5]: 22,342,240,000 pages per year.
Impressions (in-plant) All Devices: 2,750,000,000 per year
Impressions by Digital Devices: 40,678.066 per year

[1] Deposited Content Size (ESTIMATES)
a. Based on average content submission size for archived Customer Service orders in 2003.

[2] Harvested Content Size (ESTIMATES)
a. 500 GB stored for 50,000 harvested items.

[3] Converted Content Size (ESTIMATES)
a. 32 pages/publication
b. 8 Grayscale pages (14 MB/pg)
c. 4 Color pages (42 MB/pg)
d. 20 Bitonal pages (4 MB/pg)
e. 7 paragraphes /page (224 paragraphes/document)

[4] The number includes visits to web pages, retrievals of documents through WAIS, and hits to non-picture files served through Akamai

[5] Delivery from commercial sources - procured (ESTIMATES)
a. 32 pages/publication
b. 5,000 copies per order
c. 139,639 orders per year

Delivery Channels for Hard Copy or Tangible Products:

- Agency
- Rider Agency
- Sales Program
- FDLP
- ByLaw (e.g., IES, Congress, White House)

Data Location:
Internal to GPO –
Data for approximately 286,000 publications is located on 3000 distinct databases served through approximately 120 independent applications.

External to GPO –
20 sites hosted and managed on behalf of agency customers.

User Support
Information Dissemination Call Center: 3,184 incident creations per month in FY2004 and 4,797 incident creations per month in FY2005.

NOTE: Incident equals a single problem or issue, not interactions. It is assumed that there are 10 interactions per incident.

Customer Services: 20,800 customer interactions per day [6] estimated.

	Averages		
From:	**Emails**	**Calls**	**Faxes**
Customer:	6.96	13.16	3.68
Contractor	3.2	10.48	4.88
Internal GPO	6.08	4.52	1.28
To:			
Customer:	7.08	11.04	2.88
Contractor	2.76	11	3.4
Internal GPO	4.84	5.4	0.96

Number of Users

Content Originators:	8,000 unique users per day
Service Specialists (GPO staff):	600 unique users per day
Content Evaluators (GPO staff):	50 unique users per day
Service Providers	
GPO Staff:	300 unique users per day
External:	2500 unique users per day
End Users:	700,000 unique visitors (based on IP) per month
System Administration	
Operations (GPO staff):	100 unique users per day.

[6] Estimate of approximately 104 user interactions per day per Publishing Specialist with a standard deviation of .57. Estimated 200 Publishing Specialists.

4 Anticipated Situation

4.1 Order Statistics

4.1.1 In-plant

In-Plant will trend up with the implementation of modern capacity and improved processes.

4.1.1.1 25,000 orders per year

4.1.2 Procured orders

Procured orders will trend up due to better ordering processes

4.1.2.1 175,000 orders per year

4.2 Content Size and Characteristics

4.2.1 Deposited Content

Average size of the submission for printed publications may trend down due to compression techniques and the industry trend towards PDF as a submission tool. However, GPO anticipates that audio and video submissions will increase

4.2.1.1 Print specific submissions: 75 MB per order

4.2.1.2 Rich Media submissions: TBD

4.2.2 Harvested Content

Harvested Content will grow significantly before tapering off as Deposited Content submissions mitigate the need for harvesting.

4.2.2.1 Publication: 10 MB

4.2.2.2 Publications to be harvested: 139,000 unique publications per year (first 3 years)

4.2.3 Converted Content

This content is fixed and will not grow. This data will be ingested into the system over a 3-5 year period.

4.2.3.1 6.7 petabytes

4.3 Web Traffic Statistics

4.3.1 Web traffic

Web traffic will increase dramatically due to increased content and ease of use to the system (customer data reflects that current system is hard to use).

4.3.1.1 20,000,000 downloads per month

4.3.1.2 Retrievals: 770,000,000 per year [7, 8]

4.4 Content Delivery Statistics for Hard Copy Publications

4.4.1 Tangible Content

While the trend of fewer orders for large quantity press-runs will diminish, the actual number of delivered tangible publications and pages will likely increase. In addition, the FDsys will deliver content via DIP's that will be output by devices not currently managed by GPO.

4.4.1.1 Pages (procured) [9]: 5,600,000,000 pages per year.

4.4.1.2 Impressions (in-plant) non-Digital Devices: 1,000,000,000 per year

4.4.1.3 Impressions (in-plant) Digital Devices: 2,000,000,000 pages per year.

4.4.1.4 Sales Program: 4,500,000 products shipped per year.

4.4.1.5 FDLP: 4,000,000 products shipped per year.

[7] The number includes visits to web pages, retrievals of documents, and hits to non-picture files served through Akamai

[8] This number represents an approximate 75% increase from the current situation, owing to a document collection size that will more than double and a widening of the customer base due to enhanced usability.

[9] Delivery from commercial sources (procured) (ESTIMATES)

a. 32 pages/publication
b. 1,000 copies per order
c. 175,000 orders per year

4.5 Data Location

4.5.1 Future Situation: Decentralized with redundancy; exact configuration TBD.

4.6 User Support Statistics

4.6.1 Incident Creations

Incident creations will increase in the call center due to increased number of users to the system and increase in content. However, customer interactions for Customer Services should decrease since FDsys will provide more access to data and eliminate the need to contact CS specialists.

4.6.1.1 7,500 incident creations per month

4.6.1.2 5,000 customer interactions per day

4.7 Number of Users by Class

4.7.1 User Class Contact points

The total number of users will greatly increase as the modern, user friendly system and features are released. However, the number of GPO users will diminish as functions are automated (with the exception of Content Evaluators and others who will work directly with content).

4.7.1.1	Content Originators	10,000 unique users per day
4.7.1.2	Service Specialists (GPO staff)	300 unique users per day
4.7.1.3	Content Evaluators (GPO staff)	75 unique users per day
4.7.1.4	Service Providers (GPO Staff) Service Providers (External)	300 unique users per day 1000 unique users per day
4.7.1.5	End Users	2,100,000 unique visitors/month
4.7.1.6	System Admin/Ops (GPO staff)	100 unique users per day.

GPO Baseline IT Overview

October 6, 2005

GPO Systems Engineering

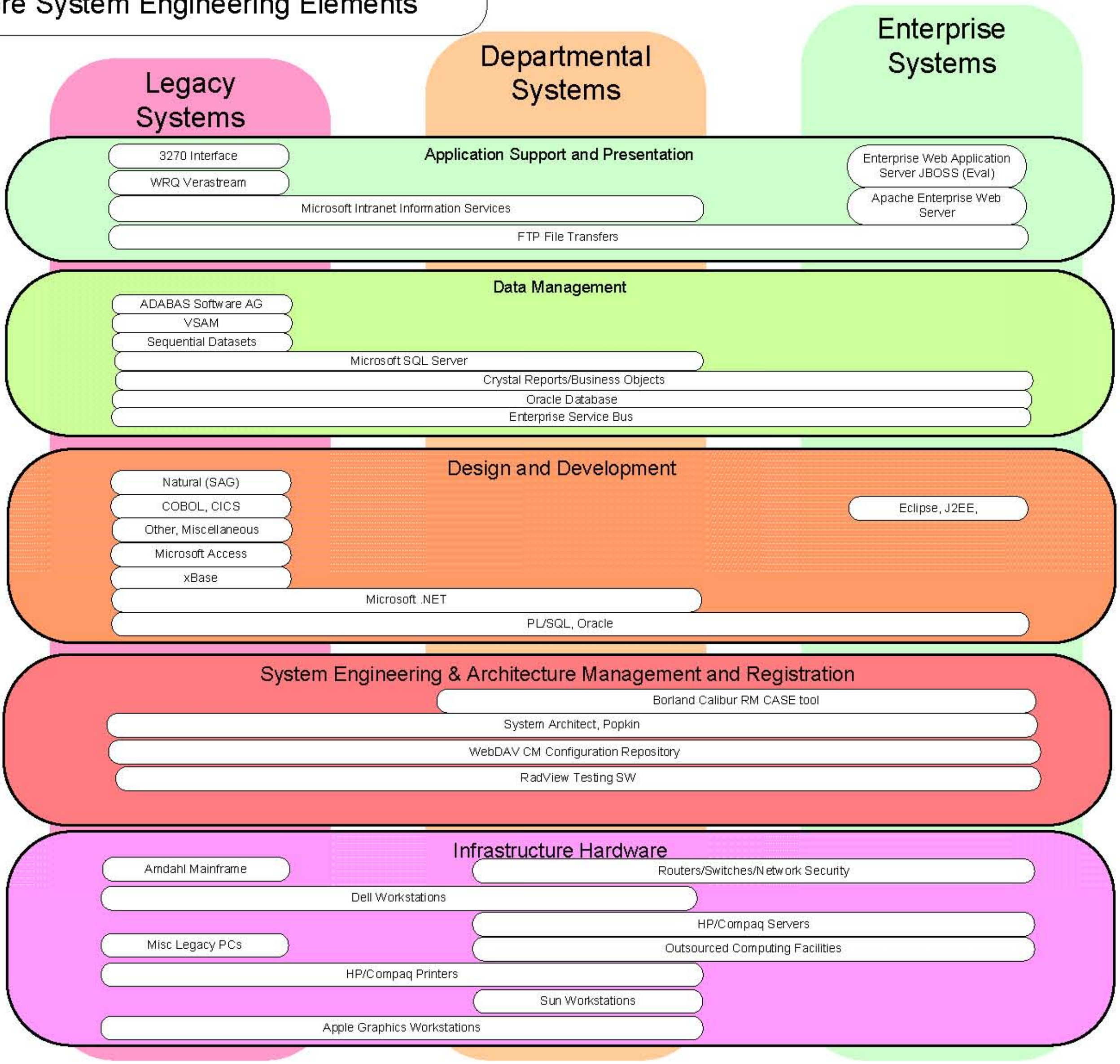
- The Government Printing Office IT infrastructure is in a process of evolving Legacy technical systems and processes to Departmental and Enterprise solutions.

- Accordingly, GPO has implemented transition projects to deliver a more streamlined and modern operation.

- GPO’s technology profile will evolve as the Enterprise Architecture matures.

- Enterprise and Departmental technologies will be relevant over the near-term to FDsys.

Core System Engineering Elements



Legacy Systems

Departmental Systems

Enterprise Systems

Application Support and Presentation

- 3270 Interface
- WRQ Verastream
- Microsoft Intranet Information Services
- FTP File Transfers
- Enterprise Web Application Server JBOSS (Eval)
- Apache Enterprise Web Server

Data Management

- ADABAS Software AG
- VSAM
- Sequential Datasets
- Microsoft SQL Server
- Crystal Reports/Business Objects
- Oracle Database
- Enterprise Service Bus

Design and Development

- Natural (SAG)
- COBOL, CICS
- Other, Miscellaneous
- Microsoft Access
- xBase
- Microsoft .NET
- PL/SQL, Oracle
- Eclipse, J2EE,

System Engineering & Architecture Management and Registration

- Borland Calibur RM CASE tool
- System Architect, Popkin
- WebDAV CM Configuration Repository
- RadView Testing SW

Infrastructure Hardware

- Amdahl Mainframe
- Routers/Switches/Network Security
- Dell Workstations
- HP/Compaq Servers
- Misc Legacy PCs
- Outsourced Computing Facilities
- HP/Compaq Printers
- Sun Workstations
- Apple Graphics Workstations

Business Oriented Architecture Principles

Principle	Directive	Enablement
Digital Information Assets	**Information is an enterprise asset**	Information and content are an enterprise asset, integral to the agency's mission. Information must be shared to enhance and accelerate decision making.
	Make information transparent	Data must be capable of being shared across the enterprise and with our partners. GPO must have a complete view of enterprise information.
Enhanced Capabilities	**Ensure security, confidentiality and privacy**	Appropriate protection in adherence with all GPO security, confidentiality and privacy policies and applicable statutesmust be in place for GPO assets throughout the architecture.
	Enable access, anytime, anywhere	System must support multiple points of access in the architecture to meet user needs.
Manage Costs	**Manage total cost of Ownership**	System application and infrastructure are expected to represent the lowest life cyle cost.
	Re-use or lease before buying, buy before building	Inventoried applications, systems, and infrastructure will be considered in the concept selection process.
	Minimize redundant efforts and data	Provide for single execution of required actions and processes in the completion GPO operations.
	Business process comes first	System functionality will be developed to support business needs.
	Reduce integration complexity	The enterprise architecture must reduce integration complexity to the greatest extent possible.

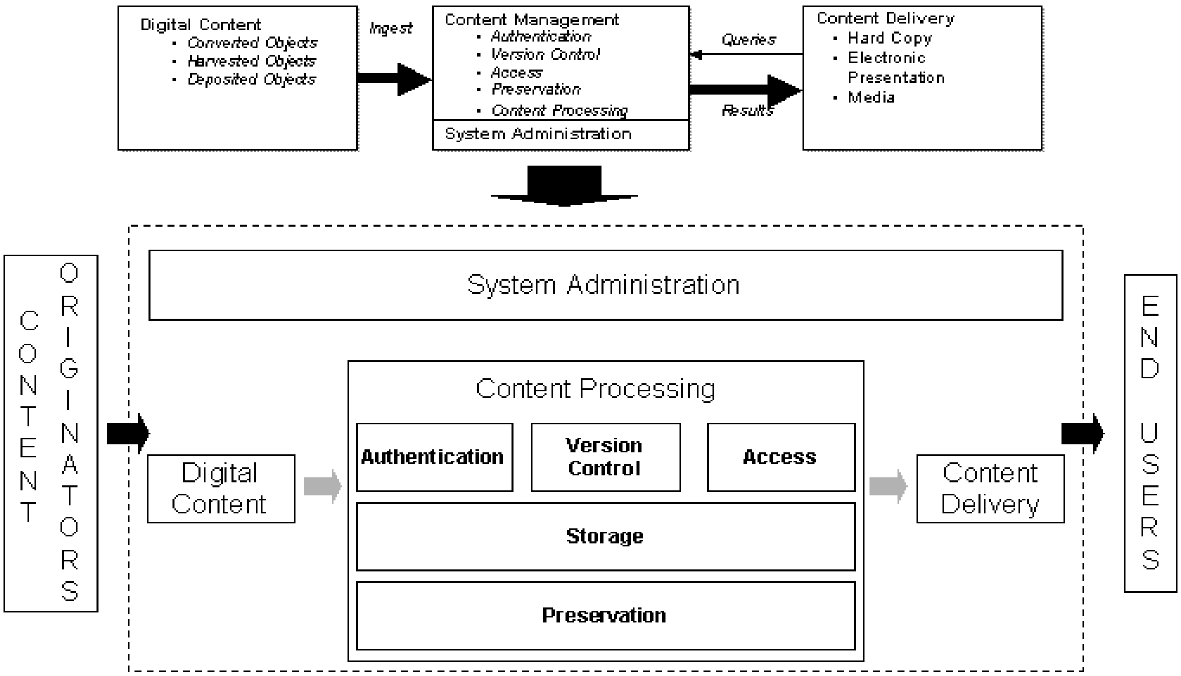
Business Continuity Architectural Principles

Principle	Directive	Enablement
Reliability	**Employ mainstream technologies**	Solutions will use industry-proven, mainstream technologies.
	Assure scalability	The underlying technology infrastructure and applications must be scalable in size, capacity, and functionality to meet changing business and technical requirements.
Interoperability	**Conform to industry standards**	Priority will be given to products adhering to industry standards and open architecture.
Availability	**Plan disaster recovery / business continuity**	Business recovery and continuity processes to meet agency needs are mandatory in GPO systems.
	Enterprise network as virtual LAN	System requires a high bandwidth GPO-wide backbone network that provides a virtual, enterprise-wide local area network.
Flexibility	**System must be policy neutral**	System requires sufficient flexibility to adapt to policy changes

Technology Oriented Architectural Principles

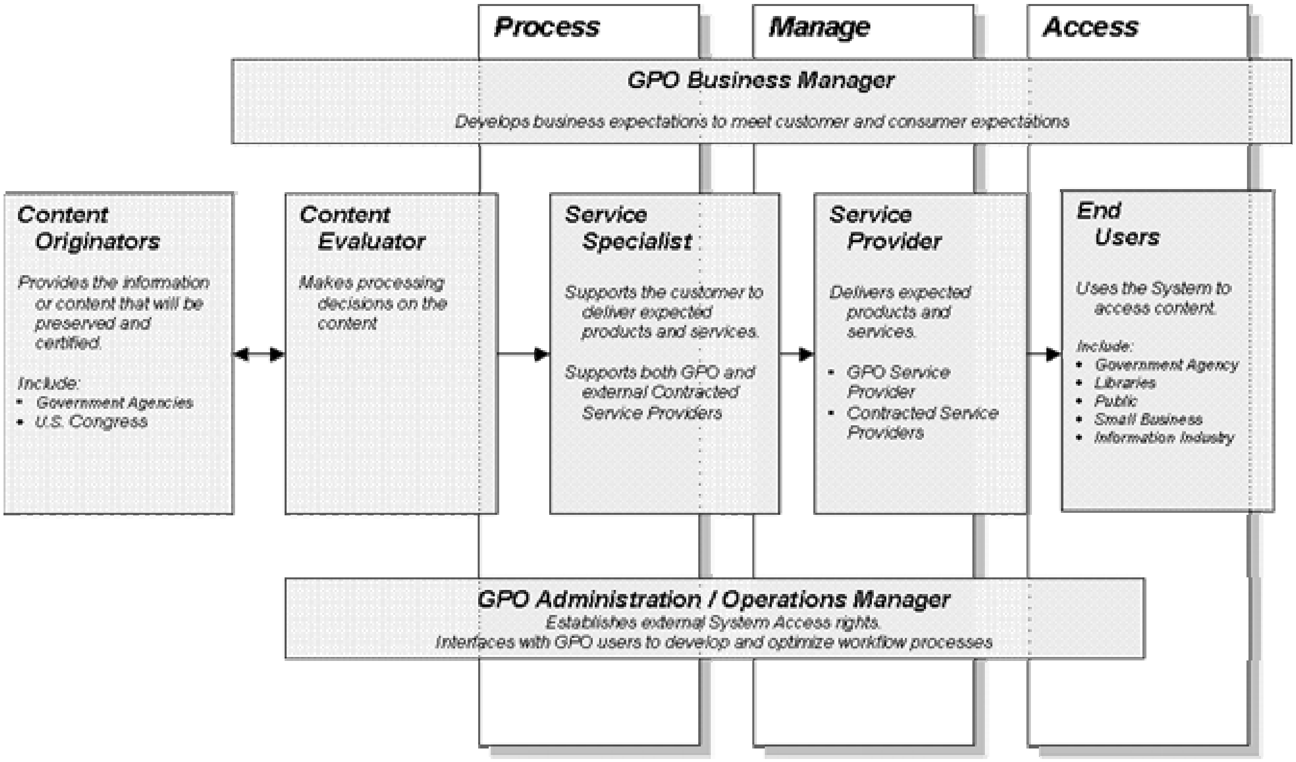
Principle	Directive	Enablement
Architecture	**Information architecture management**	System architecture must be unified and have a planned evolution that is governed across the enterprise.
	Information architecture compliance	Architecture integrity must be maintained as applications, systems and infrastructure are acquired, developed and enhanced.
Modularity of Design	**Develop with shared components using an n-tier model**	Applications, systems and infrastructure will employ reusable components across the enterprise, using an n-tier model.
	Service Oriented Architecture (SOA)	The system will be structured in a Service Oriented Architecture framework.
	Logical partitioning and boundaries	Application systems and databases should be highly partitioned with logical established boundaries that must not be violated.
	Message-based interfaces	The interfaces between separate internal and external application systems must be message-based.
Managed Infrastructure	**Logical partitioning and boundaries**	The system should consist of a small number of consistent configurations for deployment across the enterprise.
	Message-based literfaces	The system will include a standardized set of information services (e.g., email, voicemail, e-forms, user training) will be available to all GPO users.
Computing for Business Performance	**Event-driven systems**	Application systems must be driven by business events.
	Physical partitioning of processing	On-line transaction processing (OLTP) must be separated from data warehouse and other end-user computing.
	Formal software/system engineering	Consistent system engineering practices and methods based on accepted industry standards (e.g., IEEE, SEI) must be utilized.

OAIS Reference Model

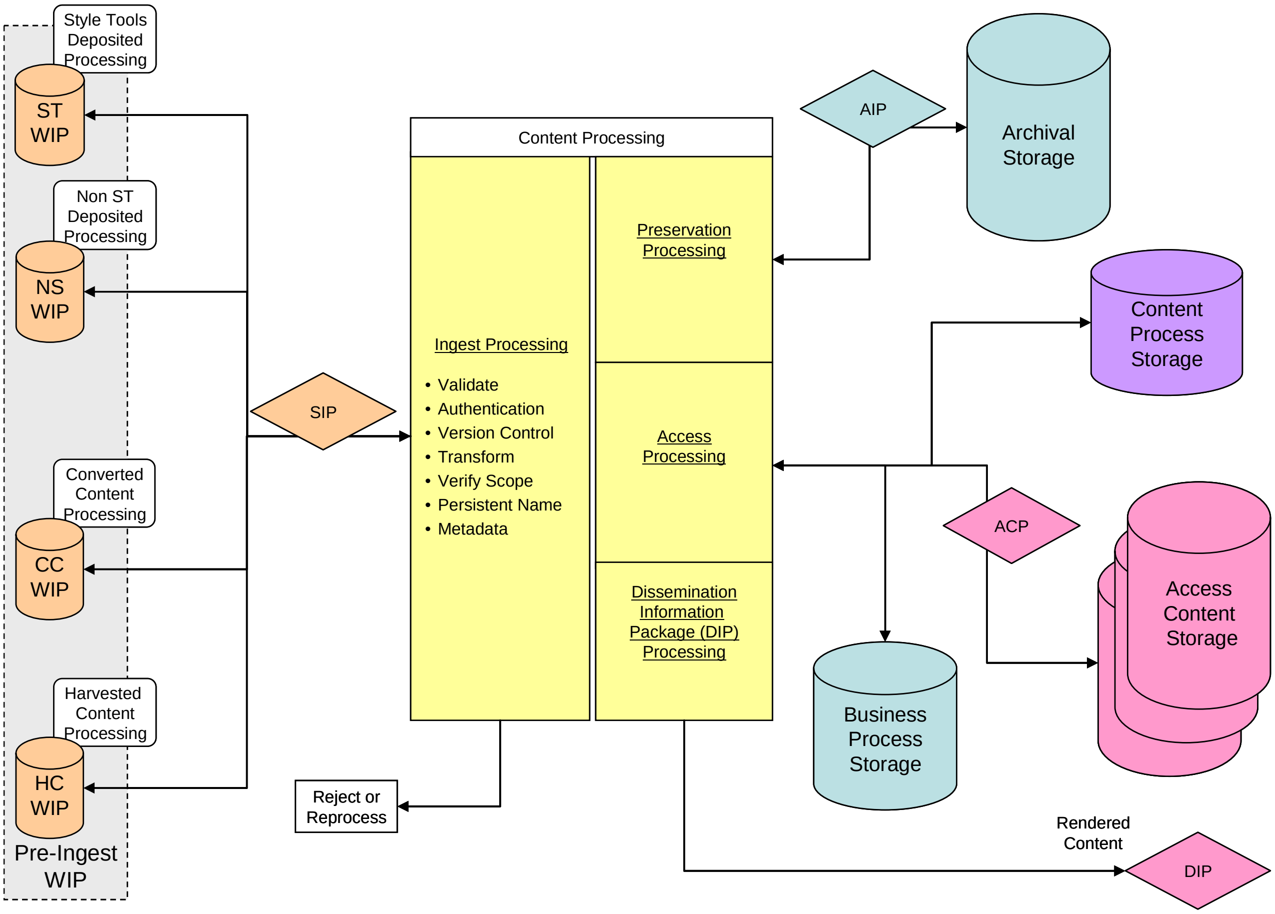


User Classes

User Classes are the fundamental groups within the broader User Categories


Process
Manage
Access
GPO Business Manager
Develops business expectations to meet customer and consumer expectations
Content Originators
Provides the information or content that will be preserved and certified.
Include:
Government Agencies
U.S. Congress
Content Evaluator
Makes processing decisions on the content
Service Specialist
Supports the customer to deliver expected products and services.
Supports both GPO and external Contracted Service Providers
Service Provider
Delivers expected products and services.
GPO Service Provider
Contracted Service Providers
End Users
Uses the System to access content.
Include:
Government Agency
Libraries
Public
Small Business
Information Industry
GPO Administration / Operations Manager
Establishes external System Access rights.
Interfaces with GPO users to develop and optimize workflow processes

Content Processing and Storage Management


Style Tools Deposited Processing
ST WIP
Non ST Deposited Processing
NS WIP
Converted Content Processing
CC WIP
Harvested Content Processing
HC WIP
Pre-Ingest WIP
SIP
Content Processing
Ingest Processing
• Validate
• Authentication
• Version Control
• Transform
• Verify Scope
• Persistent Name
• Metadata
Preservation Processing
Access Processing
Dissemination Information Package (DIP) Processing
Reject or Reprocess
AIP
Archival Storage
Content Process Storage
ACP
Access Content Storage
Business Process Storage
Rendered Content
DIP