GPO FDSYS INDUSTRY DAY
QUESTIONS AND ANSWERS

U.S. GOVERNMENT PRINTING OFFICE
KEEPING AMERICA INFORMED

10/11/05

1. Who runs New Technology & the Program Management Office (PMO)?
Ric Davis is the PMO director. The New Technology director is to be hired. Both of these organizations fall within the office of the Chief Technical Officer (CTO).

2. Will you be publishing an organization chart for the PMO?
Yes.

3. Is there a time frame for hiring the New Technology director?
Within the next 3 months.

4. Can you explain your program's funding?
As outlined in the Strategic Vision, Congress has agreed to reprogram $20 million to fund the Digital Content System. As estimated in the Strategic Vision (http://www.gpo.gov/congressional/pdfs/04strategicplan.pdf) we anticipate a $29 million spend over the next 3 years.

5. Can you describe the relationship between the master integrator and the office of the CTO?
The PMO will be the primary point of contact for the master integrator and the PMO reports to the office of the Chief Technical Officer.

6. With harvested content, how will GPO ensure content authenticity?
All content will be authenticated at the time of ingest into the system.

7. Will GPO capture all the required metadata?
Yes. The system will augment metadata throughout the content lifecycle.

8. What are your current harvesting processes?
GPO is primarily using manual processes following the GPO policy document ID73.

9. Is Microcomp or its replacement an acknowledged element of style tools?
We have developed an inventory of existing technology and applications. Each of these will be compared to requirements through our concept selection process.

10. Will the system support federated search?
Yes.

11. Is the data mining function an internal GPO tool, or can it be used by agency customers?
Data mining is an internal and external tool governed by policy.

12. Is data mining focused solely on business data or on content as well?
Data mining will be focused on business process information and metadata.

13. Have you mapped current business processes to your requirements?
The Concept of Operations document maps current processes. We expect new processes to be developed as required as the system evolves.

14. Will GPO be required to use Small Disadvantaged Business set-asides?
We do not know yet.

15. How will the PMO support the business processes?
The PMO is the point of contact for GPO's business units.

16. Will GPO define the technical architecture or will the master integrator?
GPO will have a technical architecture, but will collaborate with the master integrator to establish the final architecture.

17. Will the master integrator serve as the prime contractor and be responsible for bringing in required sub-contractors?
Yes.

18. What are you basing your schedule on?
We are working towards the completion date as outlined in GPO's Strategic Vision document.

19. Will GPO pre-select the technical components of the system?
We have developed an inventory of existing technology and applications. As we approach concept selection activities in collaboration with the master integrator, we will review the available technology and determine whether our inventory can be used.

20. After you receive responses for capability and assessment, will you provide documentation asserting that vendors who have worked with you in the past are free to work on future aspects of the system?
GPO cannot provide pre-decisional letters guaranteeing participation in an upcoming RFP.

21. How will you structure your contract to adequately compensate the master integrator?
Details are being developed.

22. Will all of the requirements be delivered within 3 years?
We anticipate having functionality for Release 1 complete in 2007.

23. If a vendor has international or commercial experience will that experience be considered?
Yes.

24. What are the key characteristics GPO is looking for in a master integrator?
We have not developed our selection criteria at this time.

25. If a company has developed a similar system for internal use, will that experience be considered?
Yes.

26. Where will development and prototyping activities take place?
We are open to on or offsite development; however, we anticipate having prototypes at GPO.

27. Is there a deployment plan and, if so, when will the beta be completed?
Yes, we have a deployment plan. Beta testing will be done in Phase 6.

28. How long will the master integrator be required to operate the beta system before it is turned over to GPO?
It depends upon the level of complexity associated with the release. GPO and the master integrator will work out the details for the release.

29. Is GPO using defined or standard terminology for development?
GPO has a SDLC document.

30. Will the SDLC document be shared?
Yes.

31. Have you have defined release procedures?
Specific release procedures will be developed for each release by the master integrator in collaboration with GPO.

32. What is the relationship between your conversion activity and the master integrator?
GPO's conversion activity is treated separately from the digital content system; however it is providing submission packages for ingest.

33. Who is responsible for interoperability between legacy systems and the FDsys? How will GPO handle legacy applications?
Legacy systems will be in use as the new system is enabled. The master integrator will work with GPO to develop the interoperability requirements.

34. Does a change management plan exist? Who is responsible?
Yes, the PMO.

35. What is the process for collaboration between the master integrator and their own technology solutions?
The master integrator and GPO will work together to select the best suited technology for the system.

36. If my company doesn't win the master integrator contract, can I still bid on FDsys solutions?
Yes.

37. Where will Q&A responses be posted from the "response" period?
FedBizOpps

38. Will the company list from Vendor Day be available?
Yes. http://www.gpo.gov/projects/pdfs/GPOIndDay_Registrationlist.pdf

39. What is the extent to which technology will drive process change?
We expect business process changes as a result of emerging technologies that are adopted for use in the system. The PMO will support the required business process changes as required.

40. Have all work products and other documentation produced by GPO been made publicly available?
Yes. – all material is on the FDsys Web page at http://www.gpo.gov/projects/fdsys.htm.

41. Are you looking at separate security products that the master integrator implements?
Yes.

42. Is GPO looking for a content management suite or just an individual software content management component?
It should be considered a content management component.

43. Who do attendees contact for questions related to FDsys and Vendor Day?
Herb Jackson at hjackson@gpo.gov

44. Will the master integrator be able to offer/suggest outsourced options?
Yes.

45. Will the master integrator be responsible for the design portion of FDsys?
Yes.

46. Will maintenance and support be included in the beta testing phase? Will it be included in the contract?
We are treating maintenance and support separately but, these details are under development.

47. Are you interested in capability statements for specific components that are included in a specific cluster?
Yes.

48. Are you planning on putting a draft RFP out?
Yes.

49. Will the presentation material be available electronically (on the web)?
Material is currently available on the FDsys Web page at http://www.gpo.gov/projects/fdsys.htm.

50. Will the technology selections emphasize "best of breed" or will integration with a single or small number of suppliers be the emphasis?
Best of breed is the emphasis.

51. Does GPO have any early views on rich media formats for the system, such as audio, Podcasts, video, etc, and how the FDsys will handle these formats?
We have been looking at these media formats and will continue to look at media formats going forward. We will emphasize open standards.

52. Does GPO plan to use .NET or J2EE technology for the system?
Both technologies have been and will continue to be looked at. It is too early to say at that level of detail whether one or both will be used. Open standards for technology and the system are important.

53. Will Disaster Recovery and backup be a consideration and accounted for in the FDsys?
Yes. Resiliency and adequate backup capability will be planned for and implemented for the system.

54. Will there be "bake-offs" between technology contenders?
The master integrator will assist GPO with the concept and technology selections. There are no plans at the moment for formal "bake-offs".

55. Is there a set time criteria defined for how long a technology must be estimated to be useful for to make it eligible for FDsys use?
There are no set time criteria for any technology. The master integrator and GPO are evaluating potential technologies against the system capabilities. GPO requires a flexible and robust integration infrastructure to enable incremental implementation of the system that should also enable updating service components with new technologies without significant re-implementation.

56. How will GPO's PMO and the master integrator work with GPO's Innovation and New Technology (INT) office regarding FDsys technology selections?
GPO's PMO and INT are in the same group. As the PMO staffing is finalized, the two groups will work together with the master integrator.

57. Who are the customers? The public or the Depository Libraries?
Depository Libraries and the general public have both been identified as system users. Additional user classes have also been identified. Please refer to the ConOps and Requirements Document for further information.

58. Are there specific system performance requirements at this time?
The specific performance requirements will be defined going forward by GPO. The master integrator will likely assist GPO in this also.

59. Has the financial model for the FDsys, in terms of which services will be charged for and what the fees charged will be, been determined?
No, that level of financial model has not been determined yet. It will be addressed as appropriate as the program progresses.

60. Will there be forums set up for small businesses to present technology and services to the master integrator and GPO?
This has not specifically been planned yet, but it is an interesting idea.

61. Is the emphasis on COTS products/software or on unique, custom built software?
COTS products/software.

62. Will FDsys replace the ILS system?
We have developed an inventory of existing technology and applications. As we approach concept selection activities in collaboration with the master integrator, we will review the available technology and determine whether our inventory can be used.

63. Does this FDsys program require or imply IT consolidation in federal agencies or GPO?
That is not required by FDsys.

64. Is there any service-oriented architecture already implemented?
SOA is an area GPO and MI will look at when architecting the system, and we want to have some SOA elements wherever possible.

65. Are the web applications, such as Portals, out of scope?
They are in scope. Where and how they are used depends on the system architecture.

66. What workflow standards have you selected?
We will evaluate available standards to meet GPO's business needs.

67. What metadata models do you require the CMS to support?
The supported metadata models are primarily driven by user needs; no particular list has been selected, but current requirement is that the CMS must support multiple metadata models.

68. Does FDsys allow interfacing with other agencies?
Yes, system integration includes internal as well as external services.

69. Do you have different requirements for different users?
No, the requirements are not categorized according to users. The user classes describe who is doing what when accessing the system.

70. Is there any ESB implementation chosen for FDsys?
The concept selection is currently underway.

71. Is FDsys going to use existing security infrastructure for authentication?
The current security infrastructure will be evaluated to see if it meets the requirement. If needed, a new security infrastructure will be put in place to fulfill FDsys requirements.

72. Does FDsys handle classified documents?
The ConOps and Requirements Document do not prohibit the handling of classified content. However, classified content is not within the scope of the FDLP.

73. Executive branch is under budget cut for IT, is that affecting your initiative?
GPO is a legislative branch agency.

74. Does the funding include hardware, software and consulting fees?
Yes.

75. Please clarify what is meant by the term unique identifier.
A unique identifier is a character string that uniquely identifies digital objects, content packages and jobs within the system.

76. Are you looking at custom content creation as a new revenue source?
Repurposing of content for specific user groups is a potential new revenue opportunity for GPO.

77. Do you abide by privacy laws and federal policies?
GPO has privacy, security and accessibility requirements and we also adhere to Federal guidelines, policies, and laws in general.

78. Will you exercise caution in terms of granularity, since granular content can be taken out of context if not encompassed with the entire document?
We are in the process of trying to define granularity in more detail, recognizing the importance of the contextual element.

79. What are you doing in relation to following the standards of standards bodies?
FDsys is standards based. Specification documentation will reference applicable standards.

80. Does GPO view itself as a leader in setting standards in the future?
GPO will participate in the development of standards.

81. We would like to see a complete inventory of all of your legacy systems that will map to FDsys.
This information will be made available to the master integrator.

82. How do you define "publication" when you use it in the Cataloging functional area?
A work of the United States Government, regardless of form or format, which is created or compiled in whole or in part at Government expense, or as required by law.

83. Do you have an established metadata standard for Cataloging?
The system shall output metadata in formats specified by the user. These formats include MARC, ONIX, and Dublin Core.

84. Do you intend to automatically create metadata?
Cataloging requirement CR1.5 in the FDsys Requirements Document Version 1.0 states that the system shall support the creation of metadata by the system (e.g. automatically create).

85. Is there a way to correlate the System Releases and Capabilities document to the Requirements Document?
The System Releases and Capabilities document presents the schedule of capabilities to be rolled out at a higher level than the Requirements Document. When the system specifications are published, they will contain release assignments for each individual system requirement.

86. Are you interested in getting feedback about the release schedule in the White Papers?
Yes.

87. Are GPO e-mails and attachments considered part of the FDsys?
User Support requirement CR1.4 in the FDsys Requirements Document Version 1.0 states that system shall provide multiple methods of contact for user assistance (e.g., web, phone, email, postal mail, chat). Emails not associated with the User Support function will not be part of the FDsys.

88. How will Access handle multiple versions?
Access will provide version information to users.

89. Is the user going to be aware of version histories?
Yes.

90. Will there be access capabilities for tangible documents?
Cataloging records for documents in the Federal Depository Library Program will be accessible through the system.

91. If the tangible document needs to be accessed, can users go to Federal Depository Libraries?
Yes.

92. Does fee-based delivery replace deposit accounts?
The system should allow for multiple methods of payment – the need to support deposit accounts is specifically mentioned in the Requirements Document.

93. Can you elaborate on content packages?
Please refer to Figure 5.5 in the FDsys ConOps 2.0 for details on all Content Packages.

94. Are Federal depository libraries included in the "library" user class?
Yes.

95. Does content access include an interface with the SuDocs sales program?
FDsys will support fee-based requests for content delivery.

96. Is GPO building a taxonomy?
While GPO is not currently in the process of building a taxonomy, we are considering all options for structuring and organizing content for access purposes.

97. Are taxonomies an integral part of Access?
We are open to leveraging innovations in knowledge management and we will evaluate each solution against the requirements.

98. You mentioned access to the data without collecting personal information. Will users be able to customize the user interface?
Yes, users will be able access information anonymously or can choose to login and create a profile that will store their preferences.

99. Will FDsys be Section 508 compliant?
GPO is committed to providing Section 508 compliant access to government information, and content delivered to users must be Section 508 compliant.

100. Do you have a more detailed description of users classes, including what features and functionalities might be of interest to each user class?
We are developing more detailed user class descriptions.

101. Will some user classes use FDsys more frequently than others?
The anticipated number of system users is listed in the System Sizing Specification document, available on the FDsys Web page at http://www.gpo.gov/projects/fdsys.htm.

102. How are you defining success?
Success is defined as meeting the requirements on time and within budget.

103.Some technologies will address multiple functional areas. Should we submit solutions multiple times?
No, just reference which solutions the technologies can be used for.

104.How does content delivery tie into the Federal Depository Library Program?
Content Delivery includes mechanisms for delivering content in a method or manner that fits the requirements of Federal Depository Libraries.

105.Is billing included in Content Delivery?
No.

106.Is the potential outsourcing of the Sales Program part of FDsys?
No.

107.Will orders go through Agency Publishing Teams for review?
Content Delivery is intended to be an automated process. When needed, GPO Service Specialists will be available to answer questions and assist with ordering content.

108.Have you conferred with users to see what they want?
We are in the process of conducting Voice of Customer (VOC) initiatives with content originators and end users.

109.Is there a separate repository (storage) for work-in progress content?
Yes.

110.Will service providers still be able to bid on work?
Yes.

111.What is being authenticated? The content or the presentation of the content?
The content is authenticated upon ingest into the system. There may be a content originator need for also authenticating the presentation of the content.

112.Where are the hard copies of material from before 1993 located? (For conversion)
FDLP legacy material is located at depository libraries across the country. We are currently working on a system to allow libraries to volunteer materials for digitization.

113.Could you please elaborate on the release schedule of content delivery?
The majority of content delivery requirements have been identified for implementation in Release 1B. Enhanced functionality will occur in Release 2. Please refer to the System Releases and Capabilities document for more information.

114.Do you perceive requirements may change over time? How about releases?
We expect that our requirements will be adjusted as we acquire additional information from users regarding expected system functionality. We have a tool to manage requirement changes. Release expectations for individual requirements may also be adjusted as we work with a master integrator to refine our priorities.

115.Are you considering subscriptions for delivery of content?
Subscriptions have been identified within the Requirements Document as a possible delivery mechanism.

116.Will the harvester be a home-grown GPO system?
GPO is currently exploring numerous options to build the harvester, and will select the solution that best meets all requirements we have stipulated. GPO is about to begin a pilot project to learn more about technologies and methodologies, the results of which will be instrumental in building the overall harvesting capability in conjunction with FDsys.

117. Is the harvester a component of FDsys?
The harvester will be a module that feeds FDsys. The harvester will be one of the tools that bring content into the system.

118. Is XML markup being considered for data capture and repurposing?
Yes.

119. How will GPO handle pilots for components of FDsys?
GPO will announce any potential pilot projects or upcoming procurements via standard procurement vehicles.

120. How do we discover RFI and proofs of concept/pilots? Where is the information available?
Information will be made available on the GPO Web site (www.gpo.gov) as well as made through traditional notification channels (e.g., FedBizOpps).

121. Does a list of the current GPO style tools exist and does it meet requirements?
GPO has developed an inventory list of current technology available within GPO. The next step after this list is completed is to map the capabilities of these technologies to the user needs and system requirements we have identified. As we approach selection activities in collaboration with the master integrator, we will review the technology and determine whether our inventory can be used.

122. What kind of content will need to be converted for Release 1 of FDsys?
GPO is beginning the conversion process with tangible printed documents, and will also move to convert audio, video, and film.

123. Will content originators be required to use style tools?
Content originators will be able to deposit content without the use of style tools.

124. How will submission tools and preservation be related and leveraged off one another?
The metadata for submission information packages must be sufficient enough so that archival information packages, access content packages, and dissemination information packages can be created from the available content and metadata.

125. How is FDsys related to the NARA ERA project since they both have similar requirements and scope?
While the current initiative at NARA have elements and concerns in common with GPO's FDsys, the overall goals and scope of collections will ultimately lead to different solutions and systems. For example, NARA's mission covers a much wider universe (Federal records) and puts emphasis on preservation of the essential evidence of the record.

GPO's mandate is long-term preservation and public access to the publications of the US Government, and thus our system will focus on issues of presentation and form in addition to preservation. Additionally, GPO's system is designed to deliver content in a variety of formats to meet the user's requirements for hard or soft copy, something that neither LC or NARA's systems are designed to provide.

126. Does GPO notify agencies before automatically harvesting documents?
GPO's Office of Information Dissemination has developed a policy on Web Harvesting (ID 73), which states that GPO will notify agencies that harvesting will take place.

127. What converted content will be considered official?
Tangible documents will be used to produce authenticated converted content packages. There will be levels of authentication associated with all content in the system, and only documents that can be traced back to the content originator will be considered "official".

128. Are you only collecting static documents?
GPO intends to locate and harvest all content within scope of its dissemination programs. However, there may be instances in which it will be more practical to create metadata records that point to content. An example of this would be a collection of statistical information that is housed in a proprietary database.

129. There will be times where there will not be enough metadata present with harvested material. How will you collect metadata on harvested content?
In these cases, GPO will harvest all metadata available with the content. As with all other content, the content will need to be cataloged and disseminated as part of the collection. Manual processes may be used to develop certain metadata elements, such as descriptive information or originating agency.

130. Are you going to harvest portions of documents?
In the context of harvesting, it is GPO's intention to collect all content that is within scope of GPO's dissemination programs.

131. Is harvesting a one time process, or a continuous project of checking for revisions and updates?
In the short term, harvesting functions will be scheduled on a case-by-case basis to ensure harvesting occurs at the appropriate frequency for all Federal agency Web sites. In the future, the need to harvest should diminish as more content is deposited into the system

132. How will converted content move from a temporary to permanent data store?
Converted content packages will be created in a format compliant with FDsys requirements. When FDsys is operational, converted content packages will be migrated into the system.

133. Are you going to sell documents?
Content will be made available free of charge through the Federal Depository Library Program. GPO will also continue to sell its products and services to end users and could possibly move into offering customized products.

134. Do you have an existing CRM system?
Yes.

135. For deposited content, will you require a minimum amount of metadata from the content originator?
Yes. Tools/modules will be present in the system that will validate and verify that content packages are compliant for ingest into the system.

136. In the RD storage requirements, STO 5.1 expresses a 100-year requirement that may be in conflict with the preservation requirements.
One of the purposes of these sessions is to identify possible discrepancies and to assist us in refining our system documentation.

137. Is writing files or documents onto microfilm an acceptable preservation solution?
Preservation solutions will be evaluated against the requirements.

138. Does GPO prefer migration, refreshment, or emulation as a preservation strategy?
GPO has not decided upon a single approach to preservation processes. We expect that the preferred process(es) will be situation-dependant.

139. What is GPO doing to capture and preserve born digital documents that are printed outside of GPO?
Capturing this content results in converted and harvested content, which after ingest will be preserved in FDsys.

140. What does "faithfully reproduce" mean?

GPO's mandate is to provide permanent public access to official versions of final, published documents. In that context, presentation and an accurate rendering of the original are vital. To faithfully reproduce an original requires an exact rendering of the content, presentation, and functionality.

141. What is GPO's position on preserving the functionality of Web-based content that is linked from harvested Web content? Should it be preserved?

At a minimum, GPO would want the metadata to indicate what the link was originally to, even if the linked content is no longer "alive." We are interested in industry's proposals and capability statements regarding this issue.

142. Must all content within FDsys be permanently preserved?

Not all content managed by FDsys will be preserved. For example, as part of its business GPO produces envelopes and such for agencies. This content does not fit the scope for preservation; e.g., final, published, official documents.

143. Does GPO have a specification or boundary statement regarding content generated "on the fly" from dynamic Web sites?

GPO has defined content in scope for permanent preservation as the final, official, published versions of U.S. Government publications. We are looking for feedback on this issue for future refinement of the RD, for example the practicality of preserving entire databases or other content sets that are the source for dynamically generated products.

144. Why is GPO harvesting?

GPO's dissemination and access missions for the Federal Depository Library Program and the cataloging program are not limited to publications produced by or through GPO; the scope of those programs is Government-wide. Therefore, GPO needs to harvest fugitive documents to meet the legal mandates of those programs.

145. Is it GPO's intent to preserve both the original and the rendered versions of content?

As shown in the Concept of Operations v.2.0, GPO intends to preserve the Archival Information Package (AIP) from which can be derived the Access Content Package (ACP), and ultimately the Dissemination Information Package (DIP) provided to the end user. There may be some cases in which it is more efficient to preserve the ACP or even a frequently-request DIP for a period of time, but we expect the permanent preservation object to be the AIP.

146. What is the extent of GPO's existing digital archive and what is your intent regarding bringing it into FDsys?

GPO has less than a terabyte of archived harvested content, plus the content presently on *GPO Access*. It is our intent to bring this legacy content forward into the FDsys preservation environment.

147. What other systems have you looked at that perform similar functions?

We are examining the preservation activities and systems of NARA (ERA), the Library of Congress (NDIIP projects), and the national libraries of Australia and the Netherlands. We are also aware of numerous efforts in the university sector to develop institutional repositories.

148. In the current environment, how do you maintain the authenticity and integrity of archived digital files?

Today, GPO puts harvested digital content into a stable environment for storage and evaluation of file structure and other characteristics.

149. Will GPO normalize content for ingest?

The nature of GPO's business means that FDsys must be capable of dealing with a wide variety of different formats coming from content originators. GPO may find it necessary to "normalize" submitted content.

150. If the rendered output is different from the original product, can the result be considered authentic or official?

If there is such a variance it should be revealed in the metadata. There will be occasions when the result of the authentication process may be to express that the content product is incomplete, but describe what is missing.

As expressed in the Preservation White Paper, the preferred outcomes 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.

151. Do you anticipate that digital preservation may result in loss of contextual information?

FDsys is focused on the practical aspects of creating a trusted digital repository that can ensure permanent public access to reliable and official digital content. As a benchmark we are examining the Audit Checklist for the Certification of Trusted Digital Repositories, Draft For Public Comment, Research Libraries Group, 2005. http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf.

152. Does agency refusal to publish through GPO result in the need for Web harvesting?

We believe the fugitive document situation arises more from lack of awareness of GPO's access and dissemination missions, especially when content is published directly to the Web without any printed product being produced.

153. Has GPO defined the file type requirements for submission?

FDsys is designed upon the OAIS model, which is not very specific in defining the ingest process. GPO has not decided that only "normalized" files in one or more approved formats may be ingested, but it is recognized this may be necessary in certain cases.

154. How does GPO decide what is in scope for preservation or dissemination.

GPO follows the intent of the content originator, the publisher, as to what is produced for public release. This is based in the definition in 44 U.S.C. 1901, and content which meets that definition is in scope for preservation.

155. What metadata is to be wrapped with the content?

This is covered in the specification and requirements for metadata. In addition to descriptive, administrative, technical, and structural metadata, the system must also record and retain metadata on all preservation processes applied to a given digital object.

156. Will all content be available for print-on-demand (POD)?

The scope of the POD activity has yet to be defined, but the FDsys must be capable of supporting POD.

157. Can you compare the GPO and NARA positions on authenticity?

GPO believes that due to mission differences the GPO and NARA views of authenticity and official status, while similar, are not identical. GPO's goal is to deliver authentic and official versions of final published documents, and often user acceptance of these hinges upon the product being an exact replica of the original. We understand NARA's primary interest to be in preserving *records* as evidence of the working of the Government.

158. Will NARA's ERA and the FDsys treat official documents or files in identical fashion?

At this time we do not have sufficiently detailed information about the workings of either system to answer that question. Interoperability with ERA and other Federal systems is an FDsys goal.