U.S. GOVERNMENT PRINTING OFFICE
KEEPING AMERICA INFORMED

Authentication

Contents:

I. BRIEFING TOPIC: Authentication

I.1 SETTING THE STAGE

Authentication is a critical function of GPO's planned Future Digital System (FDSys). As outlined in the FDSys Concept of Operations document, the authentication function will verify the authenticity of digital content within the FDSys, and certify this to users accessing the content. In order to move forward with its authentication initiatives, GPO has identified the need to develop concrete policies that address the authentication and certification of electronic Government publications. In order to keep within GPO's mission to provide permanent public access to official or authentic U.S. Government publications, GPO is currently implementing a Public Key Infrastructure (PKI) initiative to ensure the authenticity of its electronically disseminated content.

GPO recognizes that as more Government information becomes available electronically, data integrity and non-repudiation of information become more critical. The primary objective of GPO's authentication initiative is to assure users that the information made available by GPO is official and authentic and that trust relationships exist between all participants in electronic transactions. GPO's authentication initiatives will allow users to determine that the files are unchanged since GPO authenticated them, help establish a clear chain of custody for electronic documents, and provide security for and safeguard Federal Government publications that fall within scope of the National Collection of U.S. Government Publications.

I.2 NEW INFORMATION

Levels of Authentication

The provenance and fixity of an electronic document is directly related to its level of authentication. GPO will inform users about a publication's integrity and chain of custody through the designation of at least 2 different levels of authentication, "authentic" and "official." GPO defines "authentic" as content that is verified by GPO to be complete and unaltered when compared to the version received by GPO. "Official" content is content that is approved by, contributed by, or harvested from an official source in accordance with accepted program specifications. There may be instances, however, where GPO will harvest information that cannot be confirmed as official by the content originating agency. An example is a publication harvested from the Internet Archive Wayback Machine. This content will be considered authentic but not official by GPO.

Integrity Mark

The process of certification will produce an integrity mark that will include certification information and may include an emblem. Integrity marks will allow users to determine if files have been changed since GPO authenticated them, and help establish a clear chain of custody for electronic documents. Emblems may be presented to users in various ways, such as a logo used in conjunction with a digital signature. GPO will also investigate emerging technologies related to the certification and authentication of non-digital content formats (e.g., digital watermarking of GPO publications downloaded and printed by users).

Emblem

GPO may provide an emblem to notify users of the authentication status of a publication in accordance with the required approval, when feasible, of the content originator. Different content formats (e.g., audio, video, etc.) will require the use of emblems that are appropriate for each format. Users may be required to initiate additional procedures to access emblems associated with different content formats.

Look and Feel

When an emblem is visibly displayed, it should contain the official GPO authentication seal and/or official seal for the publishing agency.

Placement

When an emblem is visibly displayed, it should be placed in the same location on every document. This location should not interfere with the contents of the publication (e.g., the visible emblem should not obstruct the title of the document). The upper left hand corner is a suggested placement for the visible emblem, but additional analysis will need to be performed to ensure that this will work for all electronic publications available from GPO.

Certification Information

All integrity marks will include certification information. It is recommended that the following information be available in the certification information. This information may also be contained in a digital certificate.

- Certifying organization
- Date of the signature/certification
- Digital time stamp
- Public key value
- Hash algorithm used
- Reason for signing
- Location
- Contact information
- Name of entity that certified the publication
- Level of authentication
- Expiration date of signature / certification
- Notification of changes occurring to the document

I.3 IMPLEMENTATION

Since the completion of the operational "stand-up" of a PKI, a key generation ceremony and an external audit of operations, staff in the Chief Information Officer's organization has been working on testing and evaluating several digital signing tools using GPO's PKI. The purpose of these tests is to lead to the future application of digital signatures on *GPO Access* files. The first digitally signed documents are expected to be released soon, starting with Congressional Bills of the 109th Congress. In parallel with this testing, steps are being taken to cross-certify GPO's PKI operations with the Federal Bridge Certification Authority (FBCA). GPO has been working closely with FBCA representatives to ensure that business, administrative, and technical processes related to GPO's PKI match those of the Bridge. The FBCA is a fundamental element

of the trust infrastructure that provides the basis for intergovernmental and cross-governmental secure communications.

II. REVISED ASSUMPTIONS

II.1 PKI digital signatures and successor technologies will provide GPO with the capability to certify electronic content as authentic and official.

II.2 GPO's Authentication system will provide the capability to verify and validate that deposited, harvested, and converted content are authentic and official.

II.3 GPO's Authentication system will provide date and time verification for certified content.

II.4 GPO's Authentication system will re-authenticate the version of content that has been authenticated at earlier stages in the publishing process by GPO or Content Originators. For example, if there is a digital signature attached to a file when it comes into GPO from a publishing agency, GPO will be able to record that information and carry it forward in the provenance or in the chain of custody and provide that information to users.

III. QUESTIONS TO COUNCIL, WITH COUNCIL DISCUSSION

III.1 QUESTION: Are the assumptions in this document correct?

DISCUSSION BY COUNCIL
Council stated that an additional assumption to be considered is the direct cost of authentication to users. GPO stated that the initial files to be authenticated will be PDF files, and will provide a free plug-in if necessary to view the certification information. Council also stated the assumptions are still a work in progress, and GPO concurred.

III.2 QUESTION: As GPO works toward the implementation of its strategic vision, are we approaching the issue of authentication appropriately?

DISCUSSION BY COUNCIL
Council raised the possibility of using authentication to generate revenue, as part of the strategic vision, by contracting with other Federal agencies to help them develop their own authentication mark, which GPO could apply to documents. GPO replied that if an agency felt strongly about having its own seal, GPO would work with them to develop one. However, GPO feels that its seal would be a uniform and easily recognizable image which would make authenticity more apparent to the user. In addition, asking each agency to review and agree that every publication digitized out of the legacy collection is an official copy would create an enormous barrier and burden on the agencies.

III.3 QUESTION: When should an integrity mark be visible or invisible?

DISCUSSION BY COUNCIL
Authentication marks can be placed either on the document to show a visible representation of the signature, or be viewed only from the document's properties. Council suggested that the mark should be visible, inside the document trim.

III.4 QUESTION: To what level of granularity should GPO authenticate content?

DISCUSSION BY COUNCIL
Council stated that the ideal granularity level is the smallest usable conceivable usable level of a document, for example to the notice level or proposed rule level of the Federal Register. To start, GPO will authenticate at the document level. Council concluded that the issue of granularity is probably best addressed by a focus group that can use and test the authentication solution, taking into consideration a cost/benefit analysis for subdividing and tagging data at a granular level.

III.5 QUESTION: When authentication information is available from the publishing agency, should GPO retain and display that information in addition to GPO's own integrity mark?

DISCUSSION BY COUNCIL
Council answered with a unanimous "Yes."

III.6 QUESTION: Does Council concur with the following definitions for authentic and official content?
Authentic Content: Content that is verified by GPO to be complete and unaltered when compared to the version approved or published by the publishing agency.
Official Content: Content that falls within the scope of the National Collection of U.S. Government Publications and is approved by, contributed by, or harvested from an official source in accordance with accepted program policy and procedures.

DISCUSSION BY COUNCIL
GPO clarified that official content could come from a non-Federal source, as long as there is an official chain of custody and the creation of the document was funded by the Federal government, for example, documents harvested from a university that had a contract with a government agency to maintain their information. Council stated that although there had been discussion on the correct definitions for Authentic and Official, there was no agreement on changes to the definitions and the discussion would need to continue post-conference.

IV. QUESTIONS FROM COUNCIL ADDRESSED AT THE MEETING

IV.1 QUESTION: Users at non-depository public libraries will not be able to install a plug-in. Is there another way they can authenticate content?

RESPONSE: Users without the plug-in can view the document properties, which shows certification at the point of issuance. If the plug-in is not installed or the document is viewed from a system without an Internet connection, they will not be able to verify the certification. If verification is necessary, the user can save the file to a disk, and verify it from a suitable system.

V. AUDIENCE QUESTIONS ADDRESSED AT THE MEETING

The facilitator of the Council sessions accepted questions from the audience written on GPO-supplied cards. Sixteen of 21 questions were answered during the Council session. Those questions and their answers are summarized below. Five questions held to answer at a later date, either because of time constraints or the need for a subject matter specialist to provide more detailed answers, follow the questions answered during the session.

V.1 **QUESTION:** How often will you be notified that an item is no longer authenticated? Will this be available 24/7?

RESPONSE: The document is verified every time the document is opened, as long as the system has the necessary plug-in and Internet access. GPO is not notified if changes are made to documents that make them unauthentic.

V.2 **QUESTION:** This will be a much easier sell to my patrons if the signature is from the Clerk of the House or the Reporter of Decisions of the Comptroller General. In addition to a GPO provenance signature, I would want a bill signed by Jeff Trandahl and stamped by GPO to say this hasn't been messed with since we got it from Jeff. This will make these much more saleable to the pickier users because it's more than the Public Printer or SuDocs saying, “Trust me.”

RESPONSE: This comment was noted.

V.3 **QUESTION:** Missing from the assumptions is something to the effect that PKI digitally signed documents are the accepted standard in the end user community. Note: I do not see this happening to any great degree, reinforcing the need for tangible paper copies.

RESPONSE: Current user behavior suggests that the public is accepting the digital copy for information purposes. For those who need authenticated documents, digital signatures will establish an official chain of custody.

V.4 **QUESTION:** Would it help to clarify the concept of authentication by explicitly stating that GPO is authenticating provenance, not content?"

RESPONSE: GPO’s authentication of a document is a statement that the document is complete and unaltered since the agency publication. When GPO typesets a document, authentication is assurance that the digital copy exactly matches the copy being printed.

V.5 **QUESTION:** The mark should be visible, sort of like the depository stamp on a piece.

RESPONSE: GPO's understanding of the preference of the library community is that the mark should be visible, unless the issuing agency has specifically requested otherwise. Several audience members also requested that it be visible inside the trim, as well.

V.6 **QUESTION:** Since the signature is not applicable to electronic versions and not to any print/paper off print, signatures should always be visible.

RESPONSE: This comment was noted.

V.7 **QUESTION:** The level of granularity should extend to each format used in documents, maps, tables, et cetera.

RESPONSE: This comment was noted.

V.8 **QUESTION:** Will authentic versions of bills be available on sources other than *GPO Access*, for instance, Thomas?

RESPONSE: Thomas currently links to the PDF documents on *GPO Access* through GPO's ability to give a static URL for each of those bills. In general, the re-publisher of GPO's information will have the choice to present digitally signed documents.

V.9 **QUESTION:** Will each title or issue of a serial have a stable URL to which one can link to authenticate, if not small granularity to page, bill section, et cetera?

RESPONSE: If GPO digitally signs each article in an E-journal, each article will indeed have a static URL assigned to it.

V.10 **QUESTION:** The end user should dictate granularity of the product. The process of signature should meet the need of the end user for authentication. Also stay with your definitions, i.e. the date and time of the signature is independent of the content. Official is in the eye of the producer and end user, i.e. department may say it's official, but the court using it may not agree.

RESPONSE: This comment was noted.

V.11 **QUESTION:** Is there an official definition of official in the law?

RESPONSE: GPO is researching laws that pertain to digital signatures and the definition of "official" in that context. GPO has consulted Black's Law Dictionary, and has incorporated that definition into the definition in the assumptions.

V.12 **QUESTION:** When we receive errata, particularly for tables, we paste the corrections over the error. How will this be handled in files?

RESPONSE: These are issues that GPO is aware of internally, and GPO recognizes the need to have a policy in place to handle them. GPO will need to consider whether to

imbed the errata pages or make a correction to the file. GPO will also need to come up with a policy on how to deal with the fact that documents with changes incorporated no longer represent the typeset page that was officially issued by the publisher. The McCarthy hearings were released with more than 30 pages of errata the day they were issued. GPO released one consolidated PDF file with the errata. If a user searches Google and downloads the file that does not include the errata, they will not know those errata were issued. These are issues GPO will need to explore further as a Version Control policy is developed and implemented by GPO.

V.13 **QUESTION:** Is the key difference between authentic and official content that official content has the appropriate chain of custody; e.g., publishing agency, GPO, FDL, whereas authentic content may lack the official chain of custody? Judy confirmed this while I was writing. If so, please make that distinction for us and for all constituents.

RESPONSE: GPO plans to clarify the distinction between authentic and official by providing examples to the community as we go forward.

V.14 **QUESTION:** Does GPO intent to capture documents that have been acquired and posted on the Web by third-party sources that were not released by the agency? For example, a presidential directive obtained by the Federation of American Scientists and other classified documents.

RESPONSE: GPO will not harvest restricted documents from unofficial sources, due to our relationships with other agencies and concerns over the chain of custody. GPO will not post classified information, regardless of whether it is posted elsewhere on the Web. In addition, it is GPO's policy not to post information when the issuing agency requests that it not be distributed; this policy includes Congressional Research Service reports.

V.15 **QUESTION:** How can authentication be applied to agency databases?

RESPONSE: Right now, GPO's authentication capabilities cannot be applied to agency databases.

V.16 **QUESTION:** In reference to question six, I thought I heard Judy say during points two through five that GPO would only be authenticating its own work or contributions (e.g. typeset). Therefore, can GPO ever verify official content? Is that strictly a function of an agency? This goes to the question of GPO and meaning of content.

RESPONSE: GPO does expect to authenticate documents beyond the ones we typeset. This question directly relates to the difference between official and authentic. Documents scanned or harvested by GPO will be authenticated so that users can be sure that it has not changed since the ingest process.

VI. AUDIENCE QUESTIONS ADDRESSED AFTER THE MEETING

VI.1 **QUESTION:** What plans are there for authentication of content that does not lend itself to the PDF format?

RESPONSE: The requirements that are being developed for GPO's Future Digital System include the capability to digitally authenticate all file formats in which GPO disseminates information. In the current state, GPO's technologies can only authenticate PDF files. However, GPO will continue to monitor the industry and explore technologies for authenticating different file formats.

VI.2 **QUESTION:** What different operating systems have you tried this software on? Is the plug-in out there for Mac and Linux?

RESPONSE: GPO's future requirements for authentication stipulates that if a validation tool is required, it will work on mulitple platforms and operating systems.

VI.3 **QUESTION:** Has the GPO general council concurred with the definitions for authentic and official content?

RESPONSE: The definitions GPO has developed are very much works in progress. We have consulted GPO's General Council on these definitions, and will continue to do so as they evolve over time, and certainly before they are released as the "official" definitions.

VI.4 **QUESTION:** Does the Entrust plug-in work on non-Windows versions of Acrobat? Many libraries have non-Windows public terminals (despite the minimum technical standards) – they may be non-depositories. For example, Howard County Public Library (MD) has all Linux-based public terminals. Schools especially may have a majority Mac OS terminals. Given spyware and viruses, this may be a trend, especially for public access computers – something to keep in mind.

RESPONSE: If it is required for validation, the plug-in will work on all versions of Acrobat 5.0 or higher.

VI.5 **QUESTION:** Re: The question about digital signatures remaining a part of GPO's no-fee access commitment, it is not enough that the FDL libraries be supplied with certification equipment free of charge. It is absolutely necessary that GPO pursue certification software that will allow all users, not just FDL users, to get authenticated government information on a non-fee basis.

RESPONSE: As was stated during the session, GPO will, if necessary, provide validation tools necessary to view digital signatures at no fee to all users. The question that was asked related to additional technologies that may be used in the future, such as watermarking technologies. GPO is still in the process of defining the requirements associated with the implementation of such technologies, in conjuction with the planned implementation of the Future Digital System.