

GPO's Use of Handles



OVERVIEW

In early 1999, GPO initiated the use of PURLs to provide a persistent URL for cataloging records of depository items available on line. A review of existing technologies available at the time pointed to OCLC's PURLs – Persistent Uniform Resource Locators – as the most appropriate software available to provide this service.

In the intervening years since GPO first adopted PURLs, problems have emerged in their use including, but not limited to, scalability issues intrinsic to the PURL software which was developed with the expectation that the software would be used for local products and that far fewer PURLs would be needed than GPO subsequently required. In addition, security concerns associated with PURL technology suggest the need for migration.

A review of existing persistent identifiers led GPO's FDSys Program Management Office and its contractual partner, Harris Corporation, to select CNRI's handles as the technology for persistent identification solution for FDSys and GPO's other electronic resource management applications. The open-source nature of the CNRI handles software means it will be easier to make any special modifications GPO might require; and CNRI also actively upgrades the handles software as needs arise.

Other important features the handle system provides include:

- **Uniqueness:** Every handle is globally unique within the handle system.
- **Persistence:** Uniqueness, however, does not guarantee persistence. This is a function of administrative care. However, the handles resolution service does allow the same name to persist over changes of location, ownership, and state conditions.
- **Multiple Instances:** A single handle can refer to multiple instances of a resource, at different and possibly changing locations. Applications can take advantage of this to increase performance and reliability. For example, a network service may define multiple entry points for its service with a single handle so as to distribute the service load.
- **Multiple Attributes:** A single handle can refer to multiple attributes of a resource, including associated services, available through any method at different and possibly changing network locations. Handles can thus be used as persistent entry points into an evolving world of services associated with identified resources.
- **Extensible Namespace:** Existing local namespaces may join the handle namespace by acquiring a unique handle naming authority. This allows local namespaces to be introduced into a global context while avoiding conflict with existing

namespaces. Use of naming authorities also allows delegation of service, both resolution and administration, to a local handle service.

- Secured Name Service: The Handle System provides secured name resolution and administration. The Handle System protocol defines standard mechanisms for both client and server authentication, as well as service authorization.

GPO'S PLAN

The transition from using PURLs to using handles will be a gradual one, involving a series of tests. Ultimately, the plan includes not only the implementation of new procedures for handles, but also the migration of current PURLs to handles.

In order to complete the migration of existing PURLs, they first have to resolve correctly. GPO has identified some problematic areas:

- Broken PURLs must be identified and fixed;
- PURLs associated with publications not permanently archived must have that content harvested;
- PURLs generated, but not associated with an URL, must be resolved;
- Harvested Web sites have publications that need to be reviewed for scope determination and appropriate titles need to be cataloged and assigned a persistent identifier.

Studies will be required to identify methodologies for resolution of the problems. GPO anticipates that there be a period of time in which parallel systems of PURLs and handles will need to run.

ANTICIPATED TIMELINE

Generally speaking, the timeline for the implementation of handles is predicated on the timelines established for the Future Digital System (FDsys). Handles implementation is currently scheduled to coincide with Release 1-C of FDsys in 2008.

GPO recently conducted a limited alpha test of the handles software and a beta test predicated upon the results of the alpha test will be conducted. For the beta test, a Web page containing a number of titles with associated PURLs and examples of five handles for each title will be announced. Libraries will then be able to have a side-by-side comparison of the PURLs with the handles. Libraries will also be able to see a shorter demonstration of a method GPO has devised as a long-term strategy for resolving existing older PURLs through the handle server.

ASSUMPTIONS:

- Handles provides added security over the PURLs software.
- Handles is scalable to support a large collection of digital objects.
- Handles has more administrative support and is used by more Federal publishing agencies than PURLs.
- Handles will serve as acceptable replacements for PURLs.
- Existing PURLs in bibliographic records distributed by LSCM will continue to resolve without any action required by the library community.
- Each handle can successfully resolve to more than one location.
- Use of handles will not significantly affect document retrieval time.
- Maintenance of the handles database will be roughly on par with maintenance of the PURL database.
- While LSCM is using handles in a test environment, it is anticipated that handles will react basically the same in the production environment.
- The use of handles as a persistent identifier will not slow down cataloging production and validation reports will not decrease the response time of the servers like PURLs currently do.
- Handles will be assigned at the point of ingestion into FDsys.
- Handles will be assigned using an un-intelligent naming convention.

QUESTIONS:

- Is GPO being sufficiently cognizant of the needs and concerns of the depository community in the plan of conducting the beta test and in moving from PURLs to handles?
- What insights can Council provide in conducting the beta test?
- What experiences, if any, has council had with the use of handles?

Anatomy of a handle

Below is a screen shot of a portion of the handles beta test/demonstration page illustrating the variety of handles names used in the test. There are three parts to these sample handles.

The first part -- <http://hdl.handle.net/> -- is the domain for the handles proxy server at the Corporation for National Research Initiatives (CNRI). This is technically **not** part of the handle, but is necessary to make the handles http-compliant.

The second part is the prefix to the actual handle. The prefix is 10200, system-generated by CNRI to equate to the IP address of GPO's handles server.

The third part of the handles is the suffix. This is the part that resolves to the URL to which the handle resolves. The suffixes in the handles in the beta test exemplify some variations in the suffix including (in order):

1. Handles with the ILS number as the prefix
2. Handles using the modified form of the SuDocs number
3. Handles using a sequential number (in this case the PURL which was originally assigned to the publication)
4. Handles including descriptive information about the publication, in these cases, an acronym for the issuing agency (CDC) and file format (PDF)
5. Handles with title information including file format (HTML, PDF), issuing agency (CDC), and year of issuance

The screenshot shows a Mozilla Firefox browser window titled "Future Naming Concepts - FDLP Desktop - Mozilla Firefox". The address bar displays "http://www.fdlp.gov/handles/naming-2.html". The page content includes a heading "Future Naming Concepts" and a sub-heading "Friday, 12 October 2007". The main text states: "GPO is considering various naming structures for Handles. The following table lists the various options that initially are being explored. This list of 25 titles includes Handles of a variety of Handles names." The table lists three categories of handles:

Category	Handle	Options
1 Maternal mortality and related concepts	http://purl.access.gpo.gov/GPO/LPS71530	<ol style="list-style-type: none">1. http://hdl.handle.net/10200/0005995372. http://hdl.handle.net/10200/HE20620933/3. http://hdl.handle.net/10200/715304. http://hdl.handle.net/10200/cdc/sr03_033.pdf5. http://hdl.handle.net/10200/pdf/maternalmortalityandrelatedconcepts/2007
2 American Indian adult tobacco survey	http://purl.access.gpo.gov/GPO/LPS76206	<ol style="list-style-type: none">1. http://hdl.handle.net/10200/0005952002. http://hdl.handle.net/10200/HE20762233. http://hdl.handle.net/10200/762064. http://hdl.handle.net/10200/cdc/AmericanIndian_factsheet.htm5. http://hdl.handle.net/10200/html/americanindianadulttobaccosurvey/2006
3 Blast injuries essential facts	http://purl.access.gpo.gov/GPO/LPS74777	<ol style="list-style-type: none">1. http://hdl.handle.net/10200/0005936392. http://hdl.handle.net/10200/HE2070026123. http://hdl.handle.net/10200/747774. http://hdl.handle.net/10200/cdc/blastinjuries.pdf5. http://hdl.handle.net/10200/pdf/blastinjuriesessentialfacts/2006

The browser window also shows a search bar, a calendar, and a system clock. The system clock displays "0001-09-16 07:27:33" and "00655 07:27:33".