

25 October 2016

Overview of the Digital Object Architecture (DOA)

An Internet Society Information Paper

[Contributors](#)
Chip Sharp



Table of contents

Introduction	3
What is the Digital Object Architecture?	3
What is a Handle?	3
Handle Resolution	4
Who Runs the Global Handle System?	4
Examples of systems based on the Digital Object Architecture/Handle System	5
Standards and the Handle System	6
Policy Considerations	6
Trademarks and Service Marks	8
Resources	8

Introduction

The Digital Object Architecture (DOA) and associated Handle System[®] originated at the Corporation for National Research Initiatives (CNRI) in the early 1990's based on its work on digital libraries under contract for the Defense Advanced Research Projects Agency (DARPA).¹ One of the original motivations for its design was the need to identify and retrieve information over long periods of time (on the order of tens or hundreds of years) so persistence was a critical design requirement. At the time it was developed, the Digital Object Architecture was an attempt to shift from a view of the Internet as organized around a set of hosts and the transport to reach them to a view in which the Internet was organized around the discovery and delivery of information in the form of digital objects.

What is the Digital Object Architecture?

The Digital Object Architecture is a general architecture for a distributed information storage, location and retrieval system running over the Internet. It describes essential components for operation but allows for flexibility in how it is used to provide a service, especially in how data and metadata are represented. The fundamental components of the Digital Object Architecture include:

Digital Object

A structured record containing data, state information about the data and metadata. The digital object can contain pointers to locations where related information can be found.

Repositories

Systems where the information is stored

Identifiers/Handles

A set of identifiers, called Handles, for Digital Objects that are unique, persistent and independent of the underlying physical or logical system.

Resolution System and Registries

A system used to resolve Handles into information on the location of information and its repositories. Registries define collections of objects available in repositories.

While originally DOA allowed for different resolution systems to be defined and used², over time it has become almost exclusively tied to the Handle System.

What is a Handle?

A Handle is defined as

Prefix"/" locally unique identifier

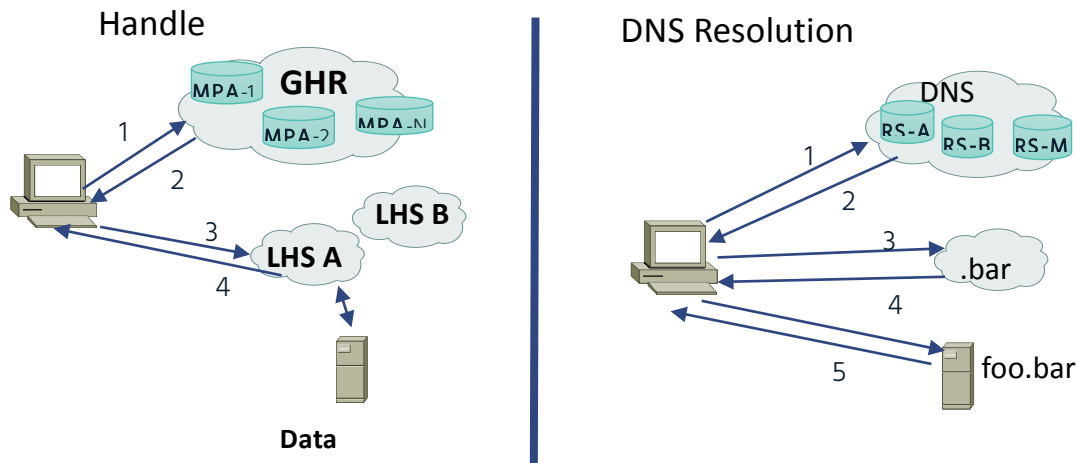
where the Prefix³ is unique within the Handle System and the "locally unique identifier" is allocated by and unique within the Prefix.

Handle prefixes are structured hierarchically where a Prefix manager can allocate sub-prefixes (separated by a period ".") to subsidiary organizations. This is similar to the Domain Name System, but the hierarchy of the Handle System is written left to right (xxx.yyy) where the hierarchy for the DNS is written right to left (yyy.xxx) where xxx is the top level. For example, the US Library of Congress has been allocated the Handle prefix "loc". The Library of Congress can then allocate the sub-prefix "natlib" for its own purposes. The complete Handle prefix would be "loc.natlib." In the DNS, the Library of Congress has the domain name of loc.gov. It has created the sub-domain for natlib whose domain name is "natlib.loc.gov."

¹ Kahn, R. & Wilensky, R., "A Framework for Distributed Digital Object Services", Int J Digit Libr (2006) 6: 115.

² Kahn, R. & Wilensky, R., op. cit.

³ Prefix was called "Naming Authority" in RFC 3651.



Handle System	DNS
1. Client sends a Handle query to the GHR for 0.NA/bar.foo	1. Client sends DNS query to a root-server for foo.bar
2. The GHR returns service information for 0.NA/bar.foo indicating LHS A manages that handle prefix.	2. The root-server returns the location of the DNS server for ".bar"
3. Client queries LHS A for bar.foo/1234	3. DNS Server A queries the .bar DNS Server for foo.bar.
4. LHS A identifies the server for bar.foo/1234. In this example, it accesses the Digital Object and returns the requested information to the Client.	4. DNS Server (B) returns requested information for foo.bar (IP address, mail server, etc.)
5. The Client then takes action on the returned information.	5. The Client then acts on the returned information (e.g. accesses a web page)

Figure 1 - Handle Resolution

Handle Resolution

The Handle System provides a method for a client to resolve a Handle into the location of a digital object. It utilizes a hierarchical service model with the Global Handle Registry (GHR) at the root and Local Handle Services (LHS) under the root. Each Local Handle Service can contain its own hierarchy of Handles Services. The GHR contains mapping information for a Handle prefix to the LHS that services the Handles for that prefix.

Figure 1 illustrates this operation for the Handle "bar.foo/1234" where the top-level prefix is "bar" and whose service information is contained in LHS A. The Handle System supports caching service information so a client doesn't always have to query the GHR similar to how a DNS resolver caches information so it doesn't always have to query the root. For comparison, Figure 1 also shows DNS name resolution for domain name foo.bar where bar is the Top-Level Domain (TLD).

Who Runs the Global Handle System?

The Global Handle Registry (GHR) is responsible for managing the root of the Handle System hierarchy allocating unique prefixes and providing a global service for mapping prefixes to the LHS for that prefix. For the first 20+ years of the Handle System, CNRI acted as the root Global Handle Registry (GHR) allocating top-level Handle prefixes. In 2015 CNRI handed over responsibility for the root of the GHR to the DONA Foundation (<http://www.dona.net>).

The International Telecommunication Union (ITU) played an important role in the creation of the DONA Foundation in 2014, working with CNRI through Memoranda of Understanding (MOUs) to develop the early plans for the transition. After its founding the DONA Foundation signed a binding MOU with the ITU in which the ITU agreed to provide secretariat support to the DONA Foundation, accept and hold the IPR and licenses on GHR technology and software

from the DONA Foundation and provide guidance to the DONA Foundation on matters of public policy. In addition to operations of the GHR, the DONA Foundation agreed to contribute its IPR to the ITU and to submit issues related to public policy to the ITU.⁴

When the DONA Foundation took over management of the GHR, it also moved to an architecture in which the GHR operation is distributed among a set of organizations called Multi-Primary Administrators (MPAs). Each MPA is allocated a top-level prefix from which it can allocate sub-prefixes and collectively with the other MPAs and the DONA Foundation carry out the functions of the GHR. Each MPA verifies and replicates all prefixes allocated by the other MPAs. Thus each MPA carries a replica of the entire Global Handle Registry.

Under this system, the DONA Foundation authorizes Multi-Primary Administrators and allocates prefixes to them. The DONA Foundation and MPAs coordinate to maintain and enhance operations of the GHR, but decision-making for policies and processes is at the DONA Board level. The current Multi-Primary Administrators as listed by the DONA Foundation can be found in Table 1.

Organization	Date	Prefix
Corporation for National Research Initiatives (CNRI) - US	3 Apr 2015	20
Coalition for Handle Services (ETIRI, CDI and CHC) – China ⁵	9 Dec 2014	86
Gesellschaft für Wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)/ePIC	9 Feb 2015	21
International DOI Foundation (IDF) -UK	1 Jan 2016	10
Communications and Information Technology Commission (CITC) - Saudi Arabia	1 Jul 2016	??

Table 1 - Current Global Handle Registry Multi-Primary Administrators

Note that while all the MPA's operate top-level prefixes, not every organization allocated a top-level prefix is an MPA. Currently available documentation on the DONA Foundation web site doesn't provide information on the role non-MPA top-level prefixes will play in the new system.

The ITU (Handle 11) was originally designated as an MPA as reflected in the minutes from the 2014 DONA Foundation Board Meeting, but is not listed as an MPA on the DONA Foundation web page. According to a keynote at the African IGF 2016 in October, the DONA Foundation signed a MPA agreement with the Republic of South Africa, but this has not yet been reflected on the DONA web site.

Examples of systems based on the Digital Object Architecture/Handle System

Several systems have based their operation on the Handle System and GHR. Each organization operating a top-level prefix can define how it uses the Handle System, e.g., semantics and syntax for its Handles and what business model it uses. It is expected that each of the MPAs will determine the operating processes and policies for use of its prefix. Examples of such systems are:

DOI® System – <http://www.doi.org>

The DOI System is managed by the International DOI Foundation (IDF), a UK-based non-profit formed in 1998 by several international publishing trade associations to support digital publishing. While it uses the GHR and Handle System, operating under the prefix 10, it defines its own syntax and semantics for Handles, metadata to be used in its system, many operational aspects of using its system and its business model. Many services operate under the DOI System such as EIDR (<http://eidr.org>) and CrossRef (<http://www.crossref.org>).

⁴ Briefing from ITU Secretary General to ITU Council in 2015 (C15/95-E)

⁵ Coalition for Handle Services (ETIRI / CDI / CHC) Consortium is jointly funded by the Institute of Electronic Science and Technology Information Technology (ETIRI) of the Ministry of Industry and Information Technology (CDI), Beijing Zhongxin Innovation and Technology Co., Ltd. (CDI), Beijing Xiandona Information Technology Co.

Persistent Identifier Consortium for eResearch (ePIC) – <http://www.pidconsortium.eu/>

The Persistent Identifier Consortium for eResearch (ePIC) provides a persistent identifier system for the European Research Community and is based on the Handle System, operating under prefix 21. It develops its own policies for allocations of its identifiers.

Standards and the Handle System

An overview of the Handle System, description of the Handle System namespace and service definitions and version 2.1 of the Handle protocol are specified in three non-consensus Informational RFCs^{6 7 8}, published in 2003. In addition, “hdl” is registered under the “info” URI scheme (info:hdl) which is defined in Informational RFC 4452⁹.

The DOI System functional components and syntax are standardized in ISO 26324:2012¹⁰ for operation under prefix 10. It imposes a specific syntax on the Handle format above and beyond the Handle System RFCs and defines additional metadata to use in the DOI System. The International DOI Foundation is the Registration Authority for ISO 26324:2012 for prefix 10. The DOI syntax was also standardized in the United States by the National Information Standards Organization (NISO) in 2005 as Z39.84-2005 (Revised in 2010)¹¹.

In 2013, the ITU-T published Recommendation X.1255¹² describing a general framework for discovery of identity management information. No protocols are standardized as part of X.1255. Although it is based on the Digital Object Architecture, X.1255 doesn't specify Handles or the Handle System protocols. There are proposals to use DOA in Study Group 20 (SG20) for the Internet of Things and for an anti-counterfeit mechanism.

Policy Considerations

The Handle System has been operated for the last 20+ years mainly for specialized use in digital libraries and research. With the GHR moving to the DONA Foundation some organizations have expressed interest in it becoming a general identifier resolution system on the Internet. However it raises a number of questions related to governance and policy.

Persistence

Persistence of identifiers and objects was one of the main goals of the Digital Object Architecture and Handle System. While the use of Handles allows for persistence, experience has shown that the main requirement for persistence is in the operation and administration of the system. For example, it doesn't matter if the system allows for persistence if the administrator forgets to update the object.

⁶ Sun, S., Lannom, L., and B. Boesch, "Handle System Overview", RFC 3650, November 2003.

⁷ Sun, S., Reilly, S., and L. Lannom, "Handle System Namespace and Service Definition", RFC 3651, November 2003.

⁸ Sun, S., Reilly, S., Lannom, L., and J. Petrone, "Handle System Protocol (ver 2.1) Specification", RFC 3652, November 2003.

⁹ Van de Sompel, H., Hammond, T., Neylon, E., and S. Weibel, "The "info" URI Scheme for Information Assets with Identifiers in Public Namespaces", RFC 4452, April 2006.

¹⁰ International Organization for Standardization (ISO), "ISO 26324:2012 Information and documentation -- Digital object identifier system", ISO Standard 26324, June 2012.

¹¹ ANSI/NISO Z39.84-2005 (R2010) *Syntax for the Digital Object Identifier*. (revised 2010)

¹² ITU-T Recommendation X.1255, Framework for discovery of identity management information, ITU-T, 2014

Name space conflicts

The Handle System allows for prefixes to consist of alphanumeric characters, e.g., “loc”, “cnri.” The Handle System faces many of the same issues that DNS registries face concerning trademarks, protected names, vanity names, etc. It is unclear how these conflicts will be addressed by the DONA Foundation and MPAs.

Governance

Transparency: Management of Internet resources requires a governance model with a high level of transparency for all of its processes and policies, e.g., how are decisions made, how is its leadership selected, how are MPAs selected and what is their agreement with the DONA Foundation, how are top-level prefixes allocated? Without such transparency it will be hard to gain the trust and confidence of the system’s users. The recent transition of IANA Stewardship illustrates the crucial role that transparency plays in the operation of global Internet identifier systems. To date, this level of transparency has not been implemented by the DONA Foundation in that very little information is available on their website (<https://www.dona.net/documents/>).

Policy Development Process (PDP): Operation of a global identifier system for the Internet requires an open, multi-stakeholder, well-defined and transparent Policy Development Process. The process must be open for review and must reflect the interests of all participants in the system as well as the Internet community. To date, based on publicly available documentation, the Policy Development Process for the GHR is opaque and limited to the Board and MPAs with no public review or consultation with the wider user community.

Protection from capture: The system must be protected from capture by any particular stakeholder group. A legitimate system must be protected from capture by any particular stakeholder group. Due to the special relationship the DONA Foundation has with the ITU (an intergovernmental, treaty organization) as evidenced in their 2014 MOU, there is concern that the system will be captured by governments and subject to geopolitical concerns rather than technical efficiency, especially in the case of a reconstitution event.

Standardization

To date most of the specifications for the Handle System have been under change control of CNRI. Specifications for global Internet identifier systems need to be developed by an open multistakeholder standards organization that follows the OpenStand principles¹³.

Economics/Business Model

While the Handle System has been used for many years in publishing and library systems, generalizing to other applications, e.g., Internet of Things, will likely generate economic concerns related to the business model of the system, especially at the Global Handle Registry. Will organizations be charged for each identifier? Will organizations that acquire a prefix be able to create unlimited sub-prefixes or will they be charged for each sub-prefix? How will these policies be developed? How will the money flow? What will be the impact on developing countries or small businesses?

Security, Stability and Resilience

Operation of a global identifier system for the Internet involves exposing both the registry of identifiers and the resolution system to a potential high degree of attack by different actors, especially as the value of the system

¹³ <https://www.open-stand.org/>

increases. The security, stability and resiliency of any such system must be understood and the system must be able to operate under severe attacks. For example, during normal operation the DNS J root server, one of the 13 root servers, sees over 6 billion queries per day¹⁴. We are not aware of any evidence to date that the GHR can handle a similar load in addition to protecting against the massive Distributed Denial of Service (DDoS) attacks seen on the Internet today.

Trademarks and Service Marks

DOI, DOI.ORG, "short DOI" are registered service marks of the International DOI Foundation, Inc.

DONA, GLOBAL HANDLE REGISTRY, HANDLE SYSTEM are registered service marks of the DONA Foundation

HANDLE.NET, HDL, HDL.NET, CNRI are registered service marks of CNRI.

HDL, HDL.NET are registered trademarks of CNRI.

Internet Society is a registered service mark of the Internet Society.

Resources

- ANSI/NISO Z39.84-2005 (R2010) *Syntax for the Digital Object Identifier*. (revised 2010)
- Corporation for National Research Initiatives, "Overview of the Digital Object Architecture", July 28, 2012 (<http://www.cnri.reston.va.us/papers/OverviewDigitalObjectArchitecture.pdf>).
- DONA Foundation. "DONA Foundation Statutes." Geneva. 2014. (https://www.dona.net/documents/public/144fc0bf2534/DONA_Foundation_Statutes.pdf)
- International Organization for Standardization (ISO), "ISO 26324:2012 Information and documentation -- Digital object identifier system", ISO Standard 26324, June 2012.
- ITU-T Recommendation X.1255, Framework for discovery of identity management information, ITU-T, 2014.
- Kahn, R. & Wilensky, R., "A Framework for Distributed Digital Object Services", *Int J Digit Libr* (2006) 6: 115.
- Paskin, Norman. "The Digital Object Identifier: From Ad Hoc to National to International." *The Critical Component: Standards in the Information Exchange Environment*, edited by Todd Carpenter, ALCTS, 2015. (https://www.doi.org/topics/150628_DOI_Case_Study_Paskin.pdf) published in
- Peter J. Denning & Robert E. Kahn, "The Long Quest for Universal Information Access", *Communications of the ACM*, Vol. 53 No. 12, Pages 34-36. (<http://cacm.acm.org/magazines/2010/12/102140-the-long-quest-for-universal-information-access/fulltext>)
- Sun, S., Lannom, L., and B. Boesch, "Handle System Overview", RFC 3650, November 2003.
- Sun, S., Reilly, S., and L. Lannom, "Handle System Namespace and Service Definition", RFC 3651, November 2003.
- Sun, S., Reilly, S., Lannom, L., and J. Petrone, "Handle System Protocol (ver 2.1) Specification", RFC 3652, November 2003.
- Van de Sompel, H., Hammond, T., Neylon, E., and S. Weibel, "The "info" URI Scheme for Information Assets with Identifiers in Public Namespaces", RFC 4452, April 2006.

This document was authored by Chip Sharp with input from Internet Society staff and experts in the Internet technical community.



This document is distributed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/) license.

¹⁴ <http://j.root-servers.org/metrics/index.html>