Managing Digital Objects

by
Robert E. Kahn
Corporation for National Research Initiatives
Reston, VA

Brussels, Belgium
September 26, 2018

What is the Internet?

- Original FNC Definition still applies:
  - Global Information System that makes use of IP (its logical extensions or follow-ons), TCP (its logical extensions or follow-ons, and other IP compatible protocols), and which supports applications based on the above.

- Overall Architecture is still intact despite increases in the underlying technology by factors of 1 – 10 Million (computation, communication and storage)

Digital Object Architecture

- The DO Architecture may be viewed as a logical extension of the Internet to simplify the task of information management.

- The concept of "digital object" forms the basis for the architecture.

- A digital object (DO) is defined as a set of bits, or a set of sequences of bits, incorporating a work or portion of a work or other information in which a party has rights or interests, or in which there is value.

- Each DO has, as an essential element, an associated unique persistent identifier.

- Assumes many different kinds of Information Systems will exist over time.

- Supports interoperability! Security!

Fundamental Properties of the DOArchitecture

- Based on the same architectural ideas embedded in the Internet's architecture, and which have sustained its evolution, the three most important characteristics being:

  - **Open Architecture** (defined protocols & interfaces)
  - **Independence** from the underlying technology
  - **Minimized Complexity** for users

- It is a <u>non-proprietary</u> architecture and is publicly available without charge.

Components of the DO Architecture

- Identifier/Resolution mechanism that allots identifiers and resolves such identifiers to "state information" about the associated digital objects (DOs) - - a resolution request yields an identifier record;

- Repositories that store DOs and enable access via their identifiers; and

- Registries that store metadata about DOs and are used for searching.

Why the DO Architecture

- Provide a uniform conceptual approach to managing information in digital form and to interact directly with that digital information

- Enables DOs & other clients to find DOs from metadata, to access them by use of identifiers

- Enables interoperability between DOs using the Digital Object Interface Protocol (DOIP)

- Security is enabled by the DOIP, but made available separately by the computational environment in which its embedded.

- "Datatypes" & Datatype registries are critical to understanding a DO by computer or otherwise; And for large data sets, more generally.

DOIP Protocol

- The Digital Object Interface Protocol (DOIP) is a simple, but powerful conceptual protocol for software applications ("clients") to interact with "services" which could be either the digital objects or the information systems that manage those digital objects.

- The DOIP Specification (version 2.0) will be conveyed by CNRI to the DONA Foundation in the coming weeks for public release – https://www.dona.net

- The DOIP enables a user (or another DO) to interact with a DO based on the use of associated identifiers
  - Each action is represented by a DO; and the interface conveys the action's identifier (ID1);
  - Each target of an action is also a DO; and the interface conveys that identifier as well (ID2);

Other Properties of a Request

- Request ID
- Client ID
- Input to be handled by the operation
- Optional information for authentication & serialization
- Basically the same for a response plus status minus request details
- Unless otherwise stated, all values (other than serialized data) are encoded in UTF-8

Basic Operations

- Hello
- Op.Create
- Op.Retrieve – some or all of a DO
- Op.Update
- Op.Delete
- Op.Search – using metadata managed by a given DOIP services
- ListOperations – supported by the given service
- Also core types and a way to describe "Extended Operations"

Framework for Discovery
ITU-T Recommendation X.1255

- Based largely on the DO Architecture, X.1255: "Framework for discovery of identity management information" was approved in September 2013.

- Focused specifically on discovery and access to information in digital form, structured as digital objects, X.1255 is applicable to operational requirements for information management more generally.

- For purposes of X.1255, a digital object is defined as a digital entity; The notion of a DO and a DE are nominally identical and may be referred to interchangeably.

- X.1255 assumes an identifier/resolution mechanism and describes a data model and interface protocol.

Some Background Reading

- Kahn, Robert E., Vinton G. Cerf, "An Open Architecture For a Digital Library System and a Plan For Its Development," The Digital Library Project Volume I: The World of Knowbots, (DRAFT) March 1988, http://hdl.handle.net/4263537/2091.

- Kahn, Robert E., Robert Wilensky, "A Framework for Distributed Digital Object Services," *International Journal on Digital Libraries*, (2006) 6(2): 115-123, https://www.doi.org/topics/2006_05_02_Kahn_Framework.pdf. (First published by the authors May 13, 1995, "A Framework for Distributed Digital Object Services", http://hdl.handle.net/4263537/5001).

- Managing Access to Digital Information, Cross-Industry Working Team, May 1997, http://www.xiwt.org/documents/ManagAccess-1.pdf.

- Denning, Peter J. and Robert E. Kahn. "The Profession of IT: the Long Quest for Universal Information Access". *Communications of the ACM*, December 2010, Vol. 53, No. 32, pp. 34-36, http://doi.org/10.1145/1859204.1859218.

- ITU Recommendation X.1255, "Framework for discovery of identity management information," was approved on September 4, 2013 (the work is based largely on CNRI's Digital Object Architecture; and Robert E. Kahn, CNRI's President, served as Editor), http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11951&lang=en

- Kahn, Robert E., "The Role of Architecture in Internet Defense", America's Cyber Future: Security and Prosperity in the Information Age, Center for a New American Security (CNAS), Volume II, Chapter XII, May 2011, http://www.cnri.reston.va.us/papers/CNAS_CyberSecurity_Kahn.pdf.

Some Background Reading (Cont'd)

- Braswell, Jefferson, Lannom, Larry, Milne, Alistair, Northey, Jim, Paskin, Norman and Traub, Ken, "Response to the Financial Stability Board's Request for an Engineering Study on the Best Approach to Managing the Structure and Issuance of Legal Entity Identifiers (LEIs)," (2012), http://doi.org/10.2139/ssrn.2197269.

- Kahn, Robert E., and Patrice A. Lyons, "Representing Value as Digital Objects," *Journal on Telecommunications & High Technology Law*, Vol. 5, Issue 1 (2006), http://www.jthtl.org/content/articles/V5I1/JTHTLv5i1_KahnLyons.PDF.

  A patent application (US 20030233570 A1), *titled* **Authenticating and using digital objects,** *and based in part on the ideas expressed in this article, was filed by CNRI. It specified that the technology may be applied in managing, inter alia, the issuance and authentication of financial instruments*). *The application was later abandoned when the claims were rejected by the U.S. PTO as covered by the now expired, CNRI Patent No. 6,135,646,* **System for Uniquely and Persistently Identifying, Managing, and Tracking Digital Objects.**

- Lyons, Patrice A. and Robert E. Kahn, "The Handle System and its Application to RFID and the Internet of Things", in *RFIDs, Near-Field Communications and Mobile Payments; A Guide for Lawyers*, edited by Sarah Jane Hughes, Cyberspace Law Committee, 2013, http://hdl.handle.net/4263537/5046.