URIs, URLs, URNs....

Web Naming and Addressing

Zhou, Xin
Course: INFSCI 2560
Instructor: M. Spring
Date: Jan. 29, 2007
URIs, URLs, URNs....

URI/URL Syntax:

scheme: scheme-specific-part

http://user:pass@example.com:992/animal/bird?species=seagull#wings

\__/   __/   \_______/   _______/   \__________/__________/\__/__/\_________/_________/\_____________/_____________/\____/____/

| protocol  | login | host | port | path | query | anchor/fragment |

URN Syntax:

<URN> ::= "urn:" <NID> ":" <NSS>

urn: isbn: 0-471-94128-X
URIs, URLs, URNs....

The Request For Comments (RFC) History of UR*s:

RFC 3987:
Internationalized Resource Identifiers (IRIs)

RFC 2718:
Guidelines for new URL Schemes

RFC 2717:
Registration Procedures for URL Scheme Names

RFC 2141:
URN Syntax

RFC 1737:
Functional Requirements for Uniform Resource Names

RFC 1630:
Universal Resource Identifiers in WWW: A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network as used in the World-Wide Web

RFC 3986:
Uniform Resource Identifier (URI): Generic Syntax

RFC 2396:
Uniform Resource Identifiers (URI): Generic Syntax

RFC 1808:
Relative Uniform Resource Locators

RFC 1738:
Uniform Resource Locators (URL)
URIs, URLs, URNs....

- Tim Berners-Lee: “A hypertext link to a document ought to be specified using the most logical name as opposed to a physical address.”
- RFC 2396: “A Uniform Resource Identifier (URI) is a compact string of characters for identifying an abstract or physical resource.”
- RFC 3987: “IRIs are meant to replace URIs in identifying resources for protocols, formats, and software components that use a UCS-based character repertoire.”
People always mistaken these three terms, however, together with the original motivation he provided, Dan Connolly, 2005, proposed his understanding of them which I agreed:

- “In information management, persistence and availability are in constant tension. This tension has led to separate technologies for Uniform Resource Names (URNs) and Uniform Resource Locators (URLs). Meanwhile, Uniform Resource Identifiers (URIs) are designed to serve as both persistent names and available locations.”

It is about persistence and availability, URN is more persistent while URL has more availability.

- Dave Crocker: “Do people use and hear others use these terms (UR{N,L,I}) with consistency?”
- Jeff Hodges: “URLs are well understood in literature, but not URNs.”
The standards of UR*s are created from the fundamental demand of information sharing based on the Internet and World Wide Web. The standards defined a set of syntax and rules of the representation of the identifiers, locators and names. They still left out some of the definitions for instance the URI to the real world or the relocation of the resource. The IETF and W3C carried out these standards to reinforce the WWW technologies. The standards are good, but not perfect; there are still so many problems to solve: the IRI, the persistence of URN, and the length of URL. Together with the development of the Web technology, more and more protocols would be involved and circumstance could changes. Keep on updating is the best way to catch up with the evolution.