INTERNET DRAFT
Requirements for Uniform Resource Names
draft-sollins-urn-00.ps,txt
Expires September 26, 1994

K. Sollins & L. Masinter
MIT/LCS & Xerox PARC
March 26, 1994

Requirements for Uniform Resource Names

ABOUT THIS DOCUMENT

This document sets out the requirements for Uniform Resource Names (URNs) within a larger Internet information architecture, which in turn is composed of additionally, Uniform Resource Characteristics (URCs), and Uniform Resource Locators (URLs). URNs are used for identification, URCs for including meta-information, and URLs for locating or finding resources. It is provided as a basis for evaluating standards for URNs.

This document is therefore intended to be issued under the "information RFC" disclaimer.

The discussions of this work have occurred on the mailing list uri@bunyip.com and at the URI Working Group sessions of the IETF. The authors may be contacted at sollins@lcs.mit.edu and masinter@parc.xerox.com.

STATUS OF THIS MEMO

Presented here are the requirements for *uniform resource names* (URNs) within the overall architecture of *Uniform Resource Identification.* In order to build applications in the most general case, the user must be able to discover and identify the information, objects, or what we will call in this architecture resources, on which the application is to operate. As the network and interconnectivity grow, the ability to make use of remote, perhaps independently managed, resources will become more and more important. This activity of discovering and utilizing resources can be broken down into those activities where one of the primary constraints is human utility and facility and those in which human involvement is small or nonexistent. Human naming must have such characteristics as being both mnemonic and short. Humans, in contrast with computers, are good at heuristic disambiguation and wide variability in structure. In order for computer and network based systems to support global naming and access to resources that have perhaps an indeterminate lifetime, the flexibility and attendant unreliability of human-friendly names should be translated into a naming infrastructure more appropriate for the underlying support system. It is this underlying support system that the Internet Information Infrastructure Architecture (IIIA) is addressing.

Within the IIIA, several sorts of information about resources are specified and divided among different sorts of structures, along functional lines. In order to access information, one must be able to discover or identify the particular information desired, determined both how and where it might be used or accessed. The partitioning of the functionality in this architecture is into *uniform resource names* (URN), *uniform resource characteristics* (URC), and *uniform resource locators* (URL). A URN identifies a resource or unit of information. It may identify, for example, intellectual content, a particular presentation of intellectual content, or whatever a name assignment authority determines is a uniquely namable entity. A URL identifies the location or a container for an instance of a resource identified by a URN. The resource identified by a URN may reside in one or more locations at any given time, may move, or may not be available at all. Of course, not all resources will move during their lifetimes, and not all resources, although identifiable and identified by a URN will be instantiated at any given time. As such a URL is identifying a place where a resource may reside, or a container, as distinct from the resource itself identified by the URN. A URC is a set of meta-level information about a resource. Examples of such meta-information are: owner, encoding, access restrictions (perhaps for particular instances), and location.

With this in mind, we can make the following statement:

The purpose or function of a URN is to provide a globally unique, persistent identifier used both for recognition and often for access to characteristics of or access to the resource.

More specifically, there are two kinds of requirements on URNs: requirements on the functional capabilities of URNs, and requirements on the way URNs are encoded in data streams and written communications.

These are the requirements for URN's functional capabilities:

- Global scope: A URN is a name with global scope which does not imply a location. It has the same meaning everywhere.

- Global uniqueness: The same URN will never be assigned to two different resources.

- Persistence: It is intended that the lifetime of a URN be permanent. That is, the URN will be globally unique forever, and may well be used as a reference to a resource well beyond the lifetime of the resource it identifies or of any naming authority involved in the assignment of its name.
- Scalability: URNs can be assigned to any resource that might conceivably be available on the network, for hundreds of years.
- Grandfathering: The scheme must permit the grandfathering of existing systems. For example, ISBN numbers, ISO public identifiers, UPC product codes and the like are naming schemes which should be allowed to be embedded within the URN system.
- Extensibility: Any scheme for URNs must permit future extensions to the scheme.
- Independence: It is solely the responsibility of a name issuing authority to determine the conditions under which it will issue a name.
- Resolution: A URN will not impede resolution (translation into a URL, q.v.). To be more specific, for URNs that have corresponding URLs, there must be some feasible mechanism to translate a URN to a URL.

In addition, the following are requirements for URNs as they are encoded:

- Single encoding: The encoding for presentation for people in clear text, electronic mail and the like is the same as the encoding in other transmissions.
- Human transcribability: For URNs to be easily transcribable by humans without error, they should be short, use a minimum of special characters, and be case insensitive. (There is no strong requirement that it be easy for a human to generate or interpret a URN; explicit human-accessible semantics of the names is not a requirement.)
- Simple comparison: Two URNs are 'the same URN' if they are spelled the same. It is a goal to make the algorithm for comparing URNs as simple as possible, with no alternative encodings, optional parts, with the possible exception that URNs should probably be case insensitive, and might have optional punctuation.
- Transport friendliness: A URN can be transported unmodified in the common Internet protocols, such as TCP, SMTP, FTP, Telnet, etc., as well as printed paper.
- Machine consumption: A URN can be parsed by a computer.

The URN specification is a description of a scheme that meets these requirements. Some design decisions are a direct result of these requirements:

- To satisfy the requirements of uniqueness and scalability, name assignment is delegated to naming authorities, who may then assign names directly or delegate that authority to sub-authorities. Uniqueness is guaranteed by requiring each naming authority to guarantee uniqueness. The name of the naming authorities themselves are persistent and globally unique, by means of a central registry.

- Naming authorities that support scalable naming are encouraged, but not required. Scalability implies that a scheme for devising names may be scalable both at its terminators as well as within the structure; e.g., in a hierarchical naming scheme, a naming authority might have an extensible mechanism for adding new sub-registries.
- Naming authorities should guarantee that at some time, somehow, there might be a mapping (or the potential for a mapping) to one or more URLs. The naming authority itself need not provide the mapping from URN to URL.
- For URNs to be transcribable and transported in mail, it is necessary to limit the character set usable in URNs, although there is not yet consensus on what the limit might be.

In assigning names, a name assignment authority must abide by the preceding constraints, as well as defining its own criteria for determining the necessity or indication of a new name assignment.