The RealName System: a Human Friendly Naming scheme

Abstract

The notion of a 'Human Friendly Naming scheme' (HFN) has been discussed in a variety of contexts, as an alternative to the use of URIs or URNs as a way of designating Internet resources (see [RFC 2276], for example).

This document introduces the RealName system, and its use as a HFN. It provides a mapping for RealNames into a URN namespace ('rn') as well as a URL scheme ('vnd.rn').

This is a preliminary draft, intended to raise the discussion of HFNs and the uses of RealNames as an initial example in creating a standard for access and resolution.

1. Introduction

It has been widely recognized that the URL syntax and structure has been unfriendly (frustrating and confusing) for non-technical Internet users, and that the URN naming system retains the 'unfriendly' behavior. For example, from RFC 2276:

In contrast to URNs, one can imagine a variety human-friendly naming (HFN) schemes supporting different suites of applications and user communities. These will need to provide mappings to URNs in tighter or looser couplings, depending on the namespace. It is
these HFNs that will be mnemonic, content-full, and perhaps mutable, to track changes in use and semantics. They may provide nicknaming and other aliasing, relative or short names, context sensitive names, descriptive names, etc. Their definition is not part of this effort, but will clearly play an important role in the long run.

URLs provide a powerful mechanism to resolve the location of resources on the Internet. For many applications, though, URLs are complex, totally unpredictable, and too lengthy to memorize.

In order to facilitate the adoption of the Internet by the broad non-technical audience, it is desirable to simplify Internet navigation via a simpler, globally available, human friendly naming system.

The RealName system has such user friendly naming as a goal. By simplifying navigation on the Internet, it aims at facilitating the adoption of the Internet by novice users. The RealName System primarily focuses on World Wide Web pages that can be associated with a trade name. Examples of trade names are brands (e.g. "Tylenol"), company names (e.g. "Apple Computer Inc"), trademarks (e.g. "Coca-Cola") and advertising slogans (e.g. Nike's "Just do it"). The RealName system offers a familiar interface to users: simple everyday words in their own language. Each year, commercial entities invest vast amount of marketing dollars in order to promote their brands and ensure that these names are known by millions of consumers all around the planet. Brand names are catchy, and mnemonic for commercial reasons and cross media advertising provides powerful means to guarantee that these names are universally known. For all these reasons, brand names are great candidates as HFNs.

Another fundamental value of a RealName is that it is by nature a name and not an address. Unlike most URLs, a RealName does not contain any information about the location of the resource that it refers to. Hence, a RealName is a more persistent name than the traditional HTTP URL of a Web page. If a user has bookmarked an HTTP URL and it changes, the resource can no longer be found. In the RealName System, HTTP URLs can change without impacting navigation. This level of indirection has clear benefits for navigating on the Internet. It also requires a new piece of infrastructure: a name resolution service. The RealName resolver is a backend service that maps RealNames into URLs.

In addition to increased robustness, name resolution services can provide benefits such as the ability to access metadata and discover resources characteristics prior to accessing it. The RealName resolution service provides simple query capabilities so that a user can search the namespace and discover new RealNames.
The notions of location independence, persistence and name resolution service are core to the concept of URN as presented in RFC 2141 [1]. The RealName system also provides a URN namespace; Section 3 of this document defines a compliant URN syntax for RealNames and proposes the RealName System as a formal URN namespace; Section 4 of this document defines a compliant URL syntax for RealName locators using the (undated) form of the RealName URN.

Human Friendly Names can be considered a new class of URI that are neither URNs nor URLs. The RealName system is proposed as an initial example of HFNs, and as the basis for future standardization of the general class.

2. The RealName system as a Human Friendly Name scheme

A "RealName" is a name registered with Centraal Corporation. Centraal Corporation owns and manages the RealName repository database and is the sole assigning authority for RealNames. Names are arbitrary strings, encoded in Unicode UTF-8. The Centraal realname resolution service offers a variety of different searching and matching mechanisms for looking up and searching the database of RealNames. The result of RealName resolution is a set of metadata about the resource, including available URLs and URNs.

3. Mapping the RealName System into a formal URN namespace

The RealName system also forms the basis of a formal URN namespace. The URN namespace consists only of the canonical registered representation of a RealName. That is, while a RealName HFN might be part of a database to be searched, the RealName URN is used for canonical lookup.

A specification template is submitted according to the guidelines defined in the IETF working draft on URN Namespace Definition Mechanisms [2].

2.1. The "rn" URN namespace Specification Template

Namespace ID:
"rn" requested.

Declared registrant of the namespace:
Nicolas Popp
nico@centraal.com

Relevant ancillary documentation:
Declaration of structure:
The identifier structure is as follows:
urn:rn:<NSS>

The <NSS> being defined as:
<NSS> ::= <prefix> "/" <realname>
<prefix> ::= (hex-escaped opaque string)
<realname> ::= (hex-escaped UTF-8 encoded Unicode string)

Conceptually, the <prefix> organizes the namespace in distinct sub
name spaces and gives the RealName URN a hierarchical
structure. The <prefix> is automatically assigned by Centraal.

The <realname> component of the NSS is the document's RealName
expressed in the URN canonical form as specified in RFC 2141
[1]. A RealName is a globally unique name that has a natural
language structure.

Identifier uniqueness considerations:
The Centraal repository database enforces the uniqueness of URNs
for all subscribed RealNames as an integrity constraint. This
guarantees that a RealName URN is unique across the entire name
space.

Centraal Corporation owns and manages the RealName repository
database and is the sole assigning authority for RealNames.

Identifier persistence considerations:
A RealName will persist in the repository beyond the life of the
Web page that it points to. Nevertheless, since commercial
entities and their brands can be replaced, it is possible on
occasion that a RealName be reassigned. For example, this would be
the case if a corporation had ceased to exist and later, a new
company was incorporated under the exact same name. In such
instance, the new corporation could legitimately subscribe and be
reassigned the defunct company's RealName.

To ensure that a RealName URN always points to the same
resource, the <prefix> for the RealName URN is changed
each time a RealName is reassigned. This guarantees that the
new RealName has a different URN than the old one.

In the first implementation of the RealName System, the URN
<prefix> is the calendar year of the subscribed RealName (note
that the proposed syntax is more general to allow future
evolution). Since RealNames are subscribed on a yearly basis, and
Centraal guarantees that a RealName cannot be reassigned for at
least one year, the RealName URN is therefore persistent.

Process of identifier assignment:
Organizations and Web site owners can subscribe to a RealName
using an online subscription service. This service is available at [https://customer.realnames.com](https://customer.realnames.com). Centraal Corporation usually charges a yearly maintenance fee for each subscribed RealName. Unlike domain names that are registered on a first-come first-served basis, Centraal exercises management and adjudication processes to ensure that a RealName is assigned to the 'appropriate' Web page. Centraal's terms and conditions require a subscriber to choose a RealName that represents 'appropriate use' for the specified Web page. If Centraal judges that a subscribed RealName is not 'appropriate', it rejects it. To assess whether or not the RealName that has been chosen by the subscriber is 'appropriate', Centraal has established a department of the company to make this determination.

'Appropriate use' should be interpreted loosely as meaning: will the Internet user community anticipate coming to the aliased Web page when using the RealName for navigation. Accordingly, common terms such as "books", "music" or "cars" are improper RealNames and cannot be subscribed because they are not unique to a single commercial Web page.

Centraal Corporation is therefore the sole assigning and managing authority for the RealName System. However, in the future, it is conceivable that the <prefix> in the RealName URN could be used as a mechanism to partition the name space and delegate some of the administrative authority to a third party.

Process for identifier resolution:
Centraal operates a form-based Web resolution service at [http://www.realnames.com](http://www.realnames.com). The RealName resolvers are built on indexing and clustering technology. Hence, they can handle large numbers of resolutions and can be distributed across the Internet. As of today, Centraal has deployed two clusters of resolvers, one on the East Coast and one on the West Coast of the United States. The current service resolves more than 5 millions RealNames a week while only operating at 20% of its current capacity. In the near future, some portions of the resolution service will be delegated to Centraal's partners all around the globe.

Centraal will follow new registration guidelines and implement the mechanisms necessary to support emerging RDS standards. For instance, Centraal will subscribe to URN.NET and will maintain a list of active resolvers as NAPTR records in the DNS.

Centraal has already prototyped an experimental URN resolver implementing the NAPTR Resolver Discovery Service as described in [RFC 2168](https://tools.ietf.org/html/rfc2168) [3].

Centraal will support THTTP and implement the N2L, N2N and N2C resolution services as described in the Internet draft "URI
Resolution Services Necessary for URN Resolutions" [4].

Rules for Lexical Equivalence:
RealNames are case insensitive. Characters with and without diacritics such as accent and vowel marks are distinguished.

To give an example (using an ASCII representation of accented characters), the RealName 'la de'pe^che du midi' is not equivalent to the Realname 'la depeche du midi'. The internal representation of a RealName is Unicode 2.1.

Hence, RealName URNs should be compared based on Unicode string equivalence as described in [5]. In particular, encoding artifacts invisible to the user should be accounted for when assessing the equivalence of two RealNames.

Conformance with URN Syntax:
There are no reserved characters in the RealName System. The reserved characters of the URN syntax will be escaped as specified in RFC 2141 [1] to ensure full conformance with that syntax. In particular, the <realname> component of the NSS is the hex-encoding of the UTF-8 for that RealName. For example, the RealName 'porsche boxster' becomes:

    urn:rn:1998/porsche%20boxster

Validation mechanism:
The RealName online subscription service available at https://customer.realnames.com provides a mechanism to check whether a RealName is already in use. Subscribers can also directly contact a name space management representative by email or telephone in order to assess what RealName is appropriate for their Web page.

Scope:
The RealName System introduced in this document does not aim to provide a RealName for every Web page on the Internet. Rather, it focuses on the subset of Web pages that can be unambiguously associated with a commercial brand name, trademark or company name. RealNames will be available in all human languages.

RealNames are globally unique and usable across the entire Internet. Therefore, the scope of a RealName URN is global.

2.2. Examples

The following are examples of URNs that a RealName resolver can resolve:

    urn:rn:1998/bambi
    urn:rn:1998/bmw%20z3
    urn:rn:1998/tylenol
2.3. Security Considerations

The primary security risk in the use of identifiers is that in some way the resource reached when following a reference will not correspond to the resource intended. The RealName system maintains a centralized scope of authority, but the reliability of the mapping depends on the security of the RealName mapping system.

3. A Proposal for the RealName System as a URL scheme

This section of the document proposes the RealName System as a new URL scheme in the vendor tree. A complete registration template is presented according to the guidelines defined in [6].

3.1 Registration Template:

URL Scheme Name:
"vnd.rn" requested.

Scope:

The RealName System focuses on the subset of Web pages that can be unambiguously associated with a commercial brand name, trademark or company name. RealNames may be in any human language. RealNames are globally unique and usable across the entire Internet; thus, the scope of a RealName URL is global.

Conformance with URL Syntax and Character encoding considerations:

A RealName URL reads

\[vnd.rn:<\text{realname}>\]

where

\[<\text{realname}> :: = (\text{hex-escaped UTF-8 encoded UNICODE string})\]

There are no reserved characters in the RealName System. The reserved characters of the URL syntax will be escaped as specified in RFC 2396 [7] to ensure full conformance with the URI Syntax. This means that the scheme specific part of the RealName URL is the hex-encoding of the UTF-8 for that RealName. For example, the RealName 'porsche boxster' becomes:

url:vnd.rn:porsche%20boxster

Security Considerations:

As with RealName URNs, the primary security risk in the use of identifiers is that in some way the resource reached when following a reference will not correspond to the resource intended. The RealName system maintains a centralized scope of authority, but the reliability of the mapping depends on the security of the RealName mapping system.
NameSpace Ownership:

Centraal Corporation owns and manages the RealName repository database and is the sole assigning authority for RealNames. The RealNames are globally unique.

Interoperability considerations:

A RealName resolution service is provided on the internet at http://www.realnames.com. The resolvers implement an HTTP/XML API that can be used by a wide variety of clients to access resources on the internet using the RealName URL.

Published specification:

This document.

Applications which use this URL scheme name:

Typical web applications can use RealName URLs as a way of locating resources by their canonical RealName without invoking a search service.

Additional information:


Contact:

Nicolas Popp
Centraal Corporation
811 Hansen Way
PO Box 50750
Palo Alto CA 94303 0750
U.S.A.
Phone: (650)846-3615
Fax: (650)858-0454
Email: nico@centraal.com

Intend usage:

COMMON

Author/Change controller:

Nicolas Popp.
Centraal Corporation
Special thanks go to Yves Arrouye and Bill Washburn from Centraal Corporation for comments on earlier drafts of this document.

5. References:


7. Authors Addresses:

Nicolas Popp
Centraal Corporation
811 Hansen Way
PO Box 50750
Palo Alto CA 94303 0750
U.S.A.
Phone: (650)846-3615
Fax: (650)858-0454
Email: nico@centraal.com

Larry Masinter
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
Phone: (650)812-4365
Fax: (650)812-4333
Email: masinter@parc.xerox.com

8. Copyright

Copyright (C) The Internet Society, 1997. All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it
or assist in its implementation may be prepared, copied, published and
distributed, in whole or in part, without restriction of any kind,
provided that the above copyright notice and this paragraph are
included on all such copies and derivative works. However, this
document itself may not be modified in any way, such as by removing
the copyright notice or references to the Internet Society or other
Internet organizations, except as needed for the purpose of developing
Internet standards in which case the procedures for copyrights defined
in the Internet Standards process must be followed, or as required to
translate it into languages other than English.

The limited permissions granted above are perpetual and will not be
revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an
"AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING
TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT
NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN
WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF
MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."