Identifiers in the Web: URIs

- ## URL: locations
  - *New York Public Library, second floor, third aisle, second shelf, third book from left*

- ## URN: location-independent names
  - `QP:475.L95;  ISBN:0-19-854529-0`

- ## URC: descriptions
  - *genre:* book*, title:* The Ecology of Vision*;*
    author: J.N.Lythgoe*; Date:* 1979*;*
    Publisher: Clarendon Press, Oxford

URL Requirements

An <u>object</u> that <u>describes</u> the <u>location</u> of a <u>resource</u>

- Global scope

- parsable

- transportable in many contexts

- extensible

- not loaded with other information

Some URL schemes

- **`http://host.dom/path`**
- **`ftp://host.dom/path`**
- **`gopher://host.dom/selector`**
- **`news:group.name`**
- **`news:article-id`**
- **`mailto:email-name@host.dom`**
- **`file:///C:/dos/path`**
- **`telnet://host.dom`**

Relative URLs

- "base" + "relative URL"
  => "absolute URL"

  base   = `http://host.dom/a/b/c/d.e`

  relative = `../../frob.htm`

  =>   `http://host.dom/a/frob.htm`

- Defines what "base" is for various contexts
- Not defined in terms of scheme

Uniform Resource Names (URN)

- name independent of location; allows for replication, migration

- separate problems of
naming authority and name assignment
resolution mechanism: finding information
about the thing named
  - location(s), metadata

Some unsolved problems with URIs

- things go away
  - Material behind URLs disappears
- **`pimples.com`**
  - vanity domains for billboard use
- Apple Computer and Apple Music
  - conflicts over short names
- **`urn:hdl:MTV/I_quit`**
  - how does authority migrate?
- **`http://www.métro.paris.fr/métro`**
  - Non-ASCII names

URLs vs URNs

- Some URLs aren't really "locators"
  - `data:, mid:, news:`
- Does the URL syntax constrain the URN syntax?
- Does the URL syntax constrain all URIs?
- Will URNs actually work?

Human Friendly Names

- Attempts to create new systems of "real names" or "go words" or …

- Used in browser instead of URL

- Companies or individuals register names

- No match will invoke a search

- Name space administrator responsible for uniqueness

The top-level domain issue

- vanity domains in .com

- .au? .com?

- Hierarchy is lost

- Trademark disputes

- attempt to add new ones politically sensitive
  - monopoly
  - fairness

Standards for MetaData and the Web

- Cataloging (Dublin Core)

- Ratings (PICs)

- Digital Signatures (proving authenticity)

- Copyright (who owns this material?)

MetaData standards

- Dublin Core and RDF

- Ratings: PICS

- Signatures, copyright and digital property rights

Dublin Core

- How to "Catalog" the web?

- 15 common resource description elements
  - title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, rights

- Expression in Resource Definition Format

- Authored using WebDAV

Platform for Internet Content Selection

- ## Self-rating:
  - content providers voluntarily label

- ## Third-party rating:
  - multiple, independent labeling services
    - Services may devise their own labeling systems
    - same content may receive different labels from different services.

- ## Ease-of-use:
  - for parents and teachers ; labels from multiple sources

Signatures, copyright

- Another kind of metadata

- Another kind of 'rating'

- object-based security requires key management