Problems URIs Don’t Solve*

* but people think they should

Larry Masinter
Xerox Corporation
Palo Alto Research Center

http://www.parc.xerox.com/masinter
http://purl.org/NET/masinter

lmphone@ibm.net, lmm@acm.org,
lar4ry@hotmail.com, masinter@ricochet.net,
gsm.masinter@byedesk.com,
masinter@stanfordalumni.org, ...
Guide to the talk

• update WWW4 talk: “problems URIs don’t solve”
• a few more modern examples
• a few new problems
• new directions for solutions
Problems URIs don’t solve

- Dangling URLs
- Vanity, commerce, & ownership
- Using name for the unnamable
Recap: The UR- story

*UR*-here, *UR*-there, *UR*-UR-Everywhere

- **URL**: Uniform Resource Locator
  - where’s the resource?
  - ubiquitous
- **URN**: Uniform Resource Name
  - location independent name
- **URC**: Uniform Resource Characteristics
  - title, author, information
- **URI**: Uniform Resource Identifier
  - any of the above
The Dangling URL problem

- The Web4Lib Problem
  "I created a list of interesting sites for my users… six months later, half the links didn’t work! What can I do?"
  "Don’t worry, the URI committee is solving this problem… Location independent names!"

- The URN solution: location-independent names:
  - Create a URN for the resource
  - Links include URNs instead of URLs
  - Tada! Problem solved!
Reasons URNs didn’t solve the problem

- Stuff moves
- Stuff goes away
- Over time, everything goes away
- Authority and trust issues
Stuff Moves (in Mysterious Ways)

- **Organizations split**
  - AT&T => bellcore, belllabs, lucent, att
  - How to find old research reports? Who maintains the old namespace?

- **Organizations rename**
  - Resource naming based on authority/organizational name
  - The old organization name is someone else’s trademark
  - The current “owner” of the name isn’t an authority for the old name

- **Organizations merge**
  - names and authority change
  - The new organization now controls the
  - URN for “NAI” (Network Associates) vs. “McAfee” for virus files

- **Grad students eventually graduate**
  - University authority or credentials don’t migrate
  - Employees leave the company
  - Organizational hierarchical names are transient
Stuff goes away

• The student isn’t interested
• The company goes out of business
• The country disappears/merges/splits
• Eventually, almost everything disappears
• URN requirement: “500 years”
  – What namespaces remain from 1499?
  – How do we resolve them?
Internet Scale Naming must scale over Time

- Almost all of the approaches to scaling don’t deal with time scale
- A simplistic solution: remember URL + timestamp
- Date URLs by date in referring material
- No help for resolution
Vanity, Commerce, and Ownership now a well-recognized problem

• Now a well-recognized problem
• Generic terms: pimples.com
• Shared trademarks
  – apple the fruit, apple music, apple computer
  – delta the airline, delta the faucet-maker
• Migration of ownership
  – MTV.com
• Approaches: WIPO, Authority & policy, new TLDs, relative names
Most approaches don’t solve the problems

- Need another level of indirection? Hacking DNS TLDs doesn’t help
- Solve only this problem, but don’t solve the other naming problems
Naming the unnamable

• The “tv:” scheme
  – Tune your TV to a particular channel
    • What’s a channel? Is Tv:5 a name? A uniform name?
    • Name a station?
    • “name” a segment of time?

• The “data:” scheme
  – Every level of indirection needs a way around it
  – Include the data instead of a pointer to the data
    – data:image/gif;base64,R0lGODdhMAAwAPAAAAAAAP///ywAAAAAMAAwAAAC8IyPqcvt3wCcDkiLC7C0qwyGHhSWpjQu5yqmC
      YsapyuvUULtONmOZtfxgFz
    – data:application/vnd.webtv.tune,channel=upn,
      date=8/16/1999,time:23:00
Are Fragments names?

- The stuff after the # in your URI
- Is it part of the URI or not?
- Part of name depends on media being pointed into
- Redirection, indirection to/from components
- Naming interacts with document model
Names for non-English

• Other Western European languages: accents
  – http://www.métro.paris.fr/métro

• Hebrew and Arabic
  – Direction marks
  – Vowel markers

• Asian languages
  – Private characters for names
  – Variations with accents for Thai, Urdu

• Even for English: capitalization creating trade names from generic terms
CNRP: Common Name Resolution

How to use “common names” for resources?

- draft-popp-cnrp-goals-00.txt
  - define a protocol
    - for browsers & search sites to use
    - for ‘name space’ providers to offer
    - simpler than ‘search’
    - more featureful than ‘string lookup’
  - Implies a service
    - reviews nature of services, revenue & motivation of participants
    - Can’t be solved independently of business & social issues
  - Please review & participate