NEEDED: A SYSTEMATIC STRUCTURING PARADIGM FOR DISTRIBUTED DATA

Jerome H. Saltzer
Library 2000
Massachusetts Institute of Technology
Links

Document A, at client

Net

Link

target

Document B, at server
Uses for Links in a Library

1. Formal paper

   Citations

2. My notes

   Things I have found

3. Search Service Index

   aardvark

3. Search Service Index

   .

   .

   .

   zulu

4. Search Protocol

   aardvark
LINKS

Intellectually generated

Scholarly paper with citations
Survey paper with lots of citations
Bibliographical book
End of year journal index
Library card catalog
List of an author's works
New York Times list of name appearances
Reader's Guide, SCI, SSCI, RILA,

File containing the result of a search request
List of papers consulted by Einstein

Mechanically generated
Why it is Hard: Persistence

Over a long-enough time,

The storage service will
upgrade,
relocate,
merge, and
reorganize.

The document will
be updated or superseded,
be withdrawn,
be lost and restored,
become available from a
competitive service, and
be indexed in many places
Why it is Hard: Deduplication
Why it is Hard: Forwarding

Client

Storage Service A

Storage Service B
Who is Thinking about it:

Ted Nelson  Xanadu
Brewster Kahle  WAIS
Tim Berners-Lee  World-Wide Web
Cliff Lynch  CNI
Conclusion

• There is an opportunity here for a contribution from the systems community.