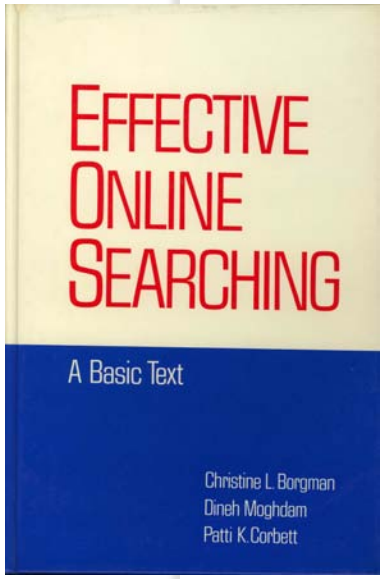
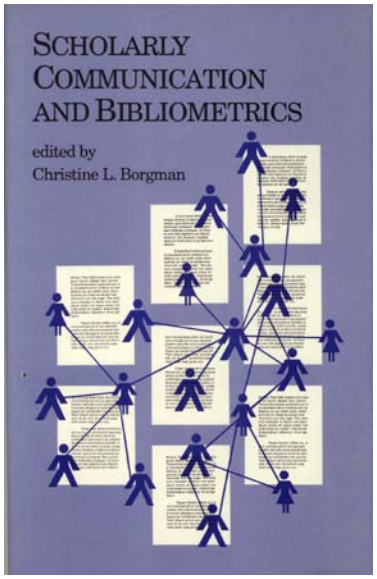
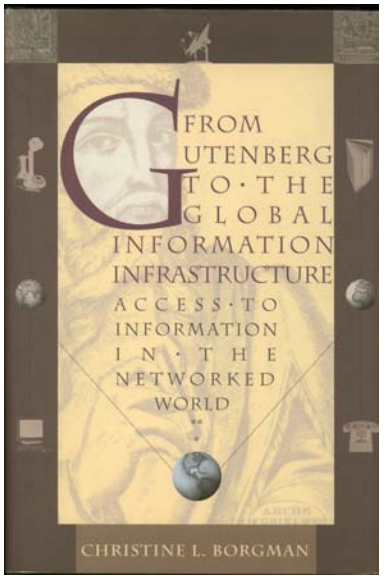
This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States of America License.

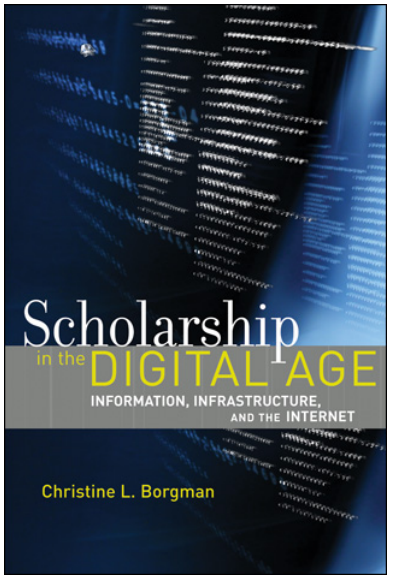
Information, Infrastructure, and the Internet

Reflections on Three Decades in Internet Time









Christine Borgman, Paul Evan Peters Award

Coalition for Networked Information Meeting, San Diego, April 4, 2011

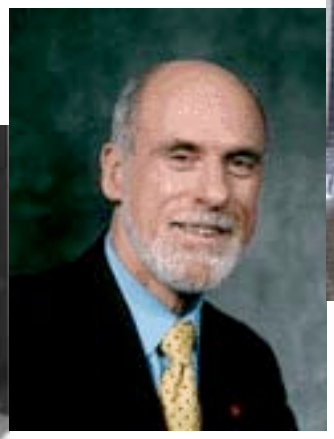
Information, Infrastructure, Internet

- From where did we come?
- Where are we now?
- Where might we go from here?

Kyoto, 2009

Information, Infrastructure, Internet

- From where did we come?



Tim Berners-Lee: Award 2000

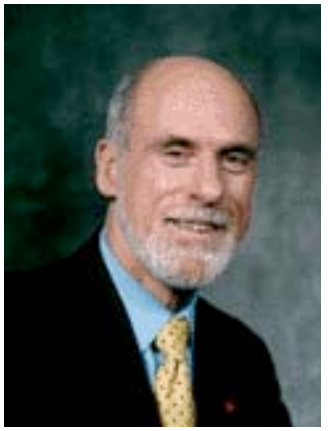
- World-Wide Web, 1991
  - HTTP
  - HTML
  - URL
- Semantic Web, 1999
- Web science, 2006

http://www.educause.edu/Professional+Development/PaulEvanPetersAwardWinnersspan/PaulEvanPeters2000AwardWinner/1516

http://home.messiah.edu/~ar1314/definitions.html

Vinton Cerf: Award 2002

- Telecommunications Protocol / Internet protocol -TCP/IP (with Robert Kahn)
- Internet Society
- ICANN
- Interplanetary Internet



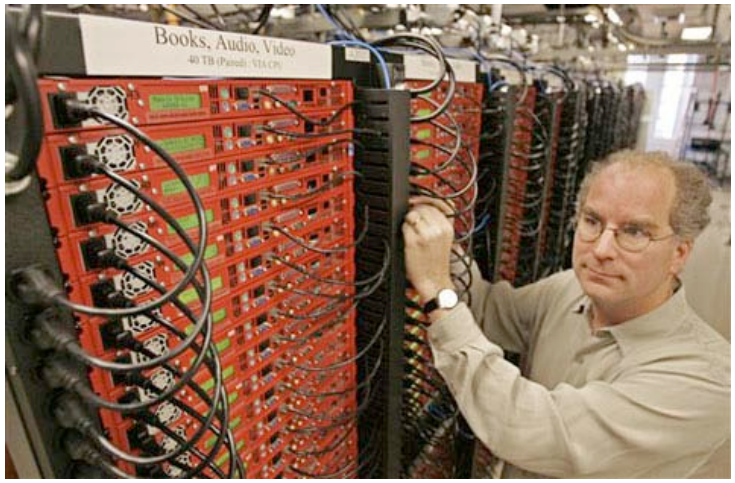
http://www.educause.edu/Professional+Development/PaulEvanPetersAwardWinnersspan/PaulEvanPeters2002AwardWinner/1515

http://en.wikipedia.org/wiki/Interplanetary_Internet

Brewster Kahle: Award 2004

- Internet Archive, 1996-
  - Web archiving
  - Book scanning
  - Contributed content
  - Personal digital archiving…

http://www.educause.edu/Professional+Development/PaulEvanPetersAwardWinnersspan/PaulEvanPeters2004AwardWinner/1514

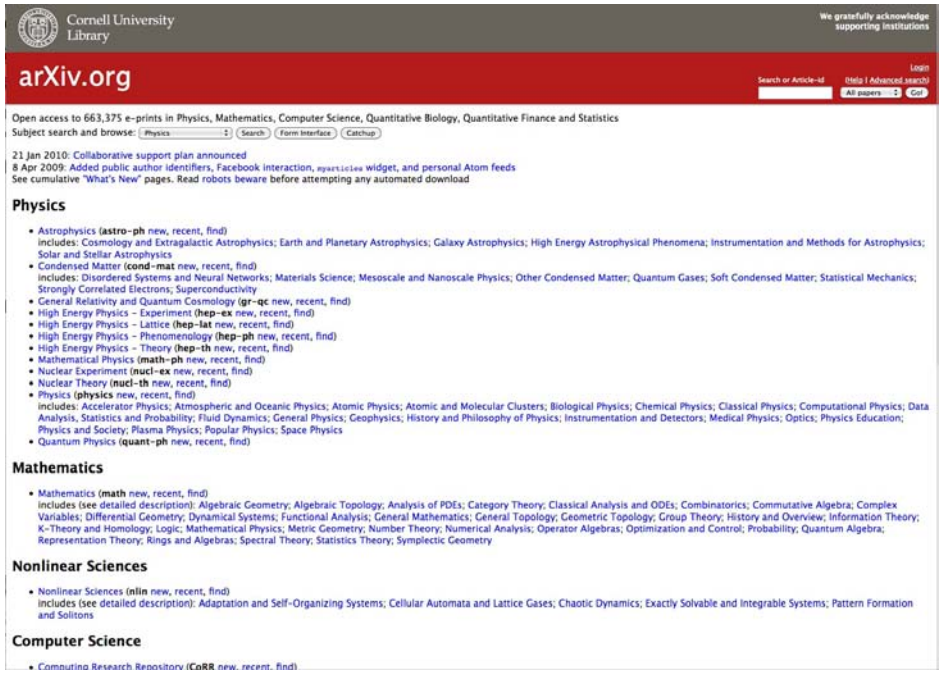


http://www.escapefromberkeley.com/race-info/advisory-board/brewster-kahle/

Paul Ginsparg: Award, 2006

- arXiv, 1991-
  - Preprint distribution
  - Open access publishing
  - Institutional repositories

http://www.educause.edu/Professional+Development/PaulEvanPetersAwardWinnersspan/PaulEvanPeters2006AwardWinner/10177

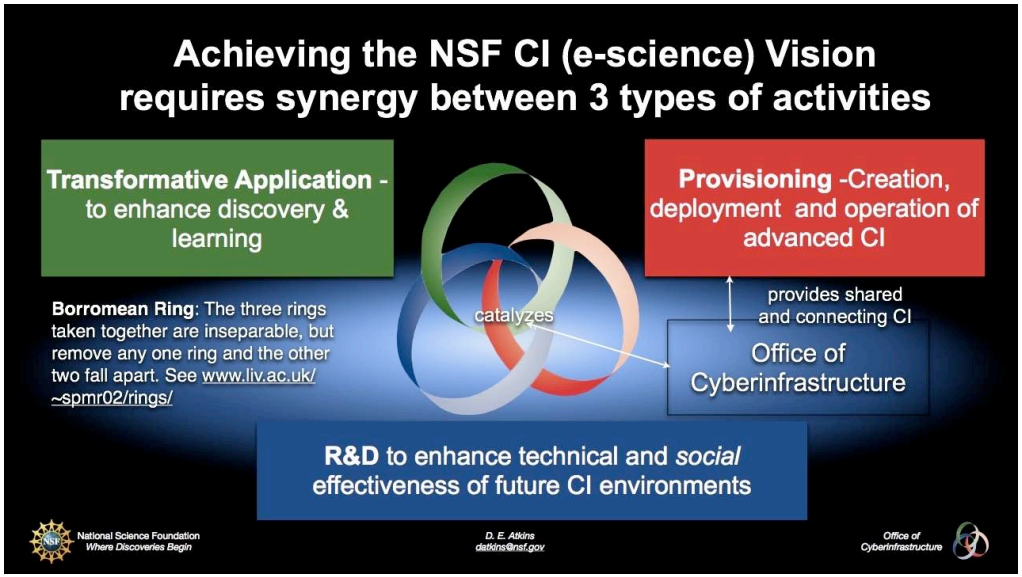


Arxiv.org homepage pi day 2011

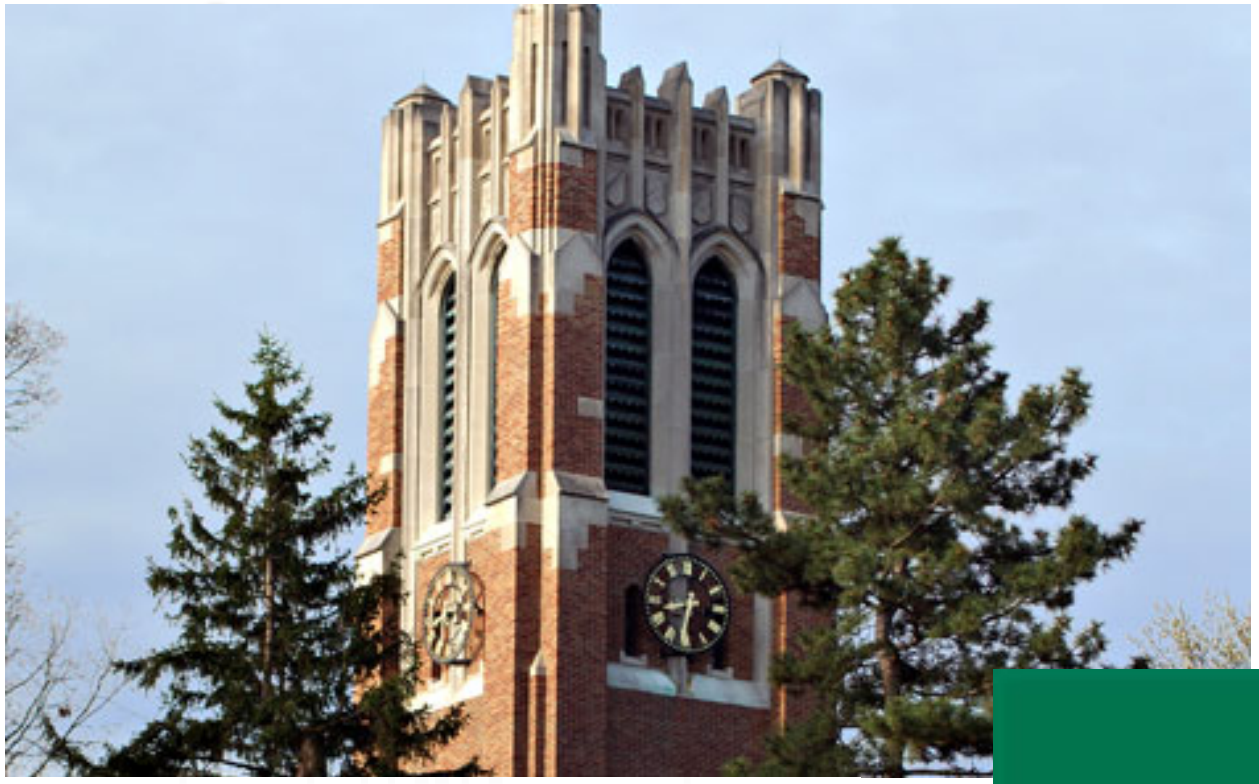
Daniel Atkins: Award 2008

- School of Information, U of Michigan
- Cyberinfrastructure @ NSF
  - Blue Ribbon Panel
  - Office of Cyberinfrastructure

http://www.educause.edu/About+EDUCAUSE/PressReleases/NSFCyberinfrastructureDirector/17312



The librarian's daughter from Detroit



π

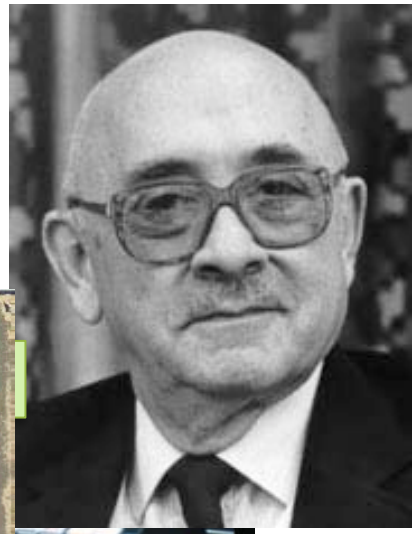
$$\int_a^b f(x)\,dx = F(b) - F(a)$$

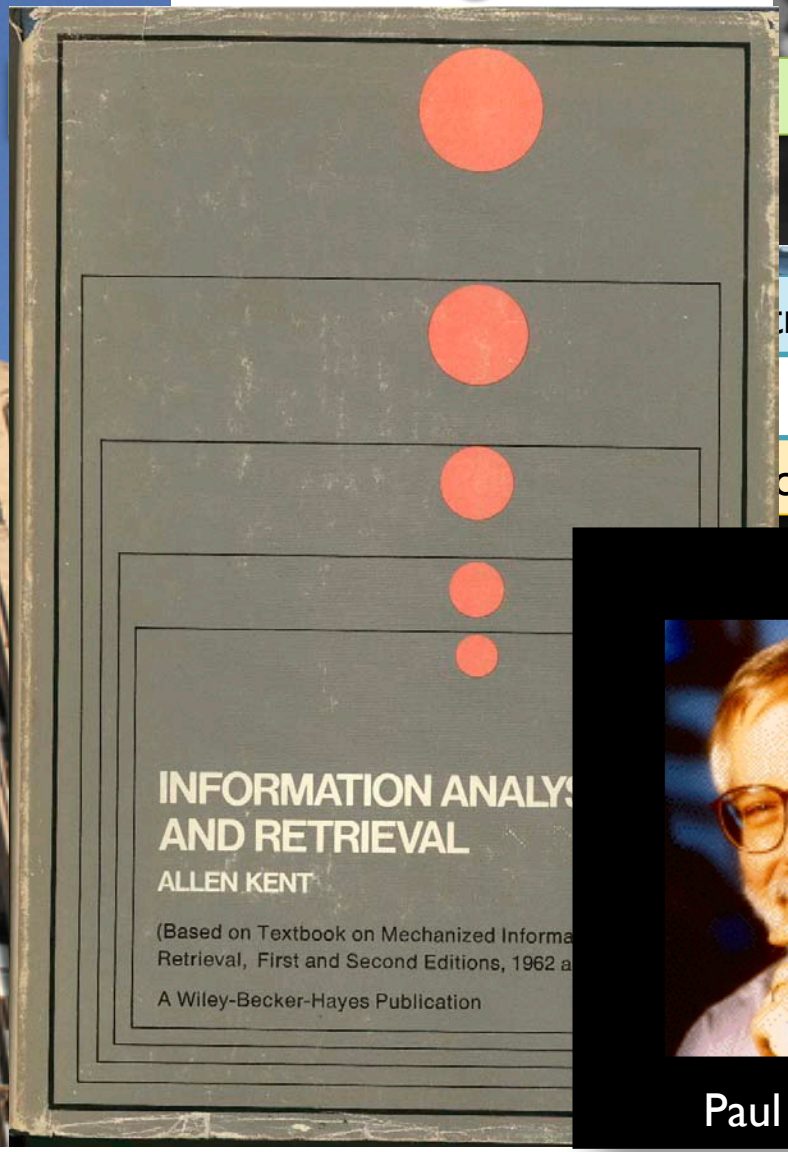
http://www.math.uga.edu/~dkrashen/courses/2200Fall2009/index.html

MICHIGAN STATE™

Allen Kent

University of Pittsburgh

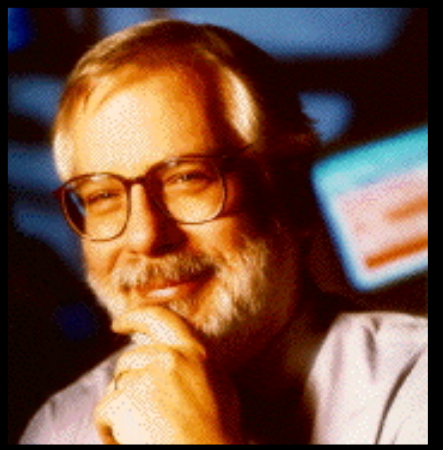




rieval System)

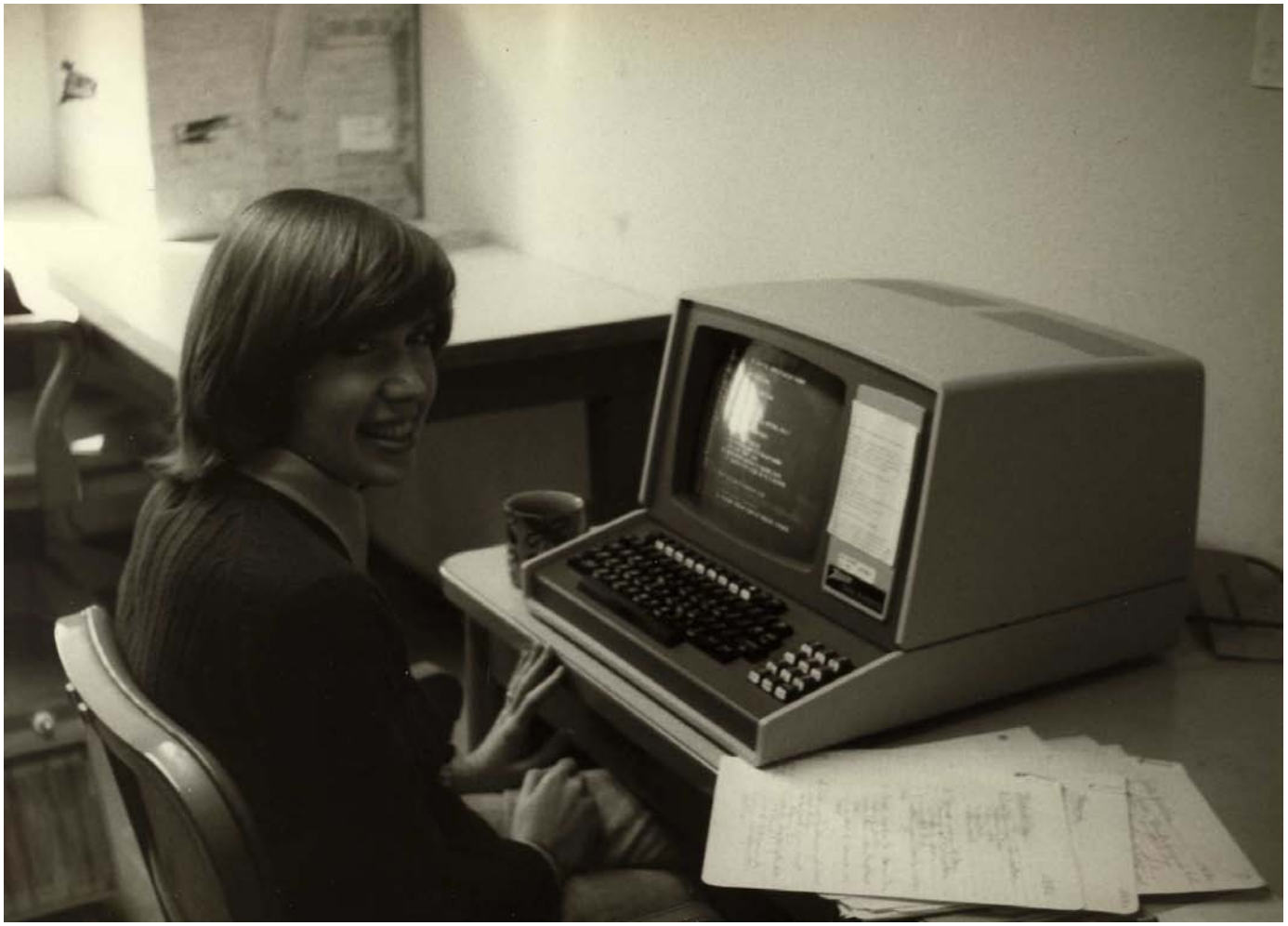
http://www.sis.pitt.edu/~mbsclass/hall_of_fame/kent.html

ormation Bank



Paul Evan Peters

Detroit girl goes to Dallas



Report Number: OCLC/OPR/RR-83/7
Date: 1983 August 31

Council on Library and Information Resources

RLG

Research Report Prepared for OCLC

on

OCLC

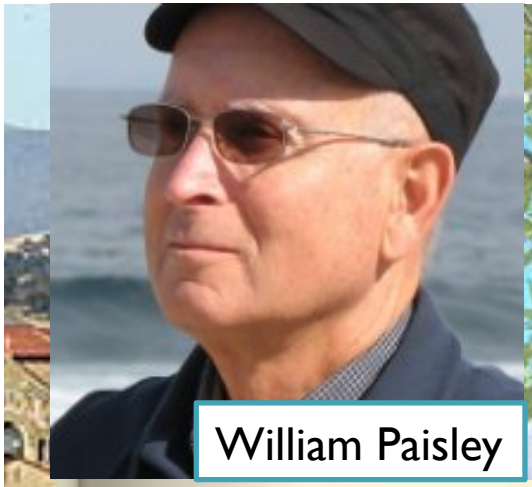
END USER BEHAVIOR ON THE OHIO STATE UNIVERSITY LIBRARIES' ONLINE CATALOG: A COMPUTER MONITORING STUDY

http://people.ischool.berkeley.edu/~clifford/

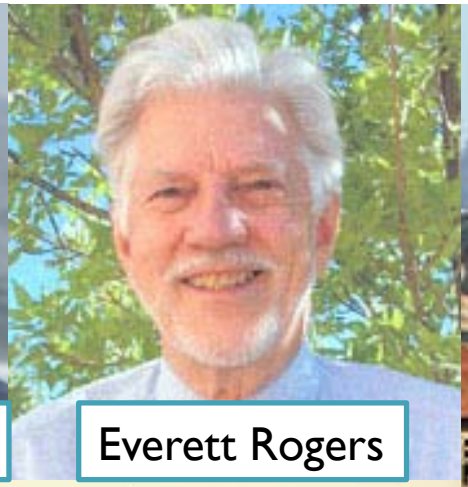
by

Christine L. Borgman

OCLC Online Computer Library Center, Inc.
Office of Research
6565 Frantz Road
Dublin, OH 43017-0702



William Paisley



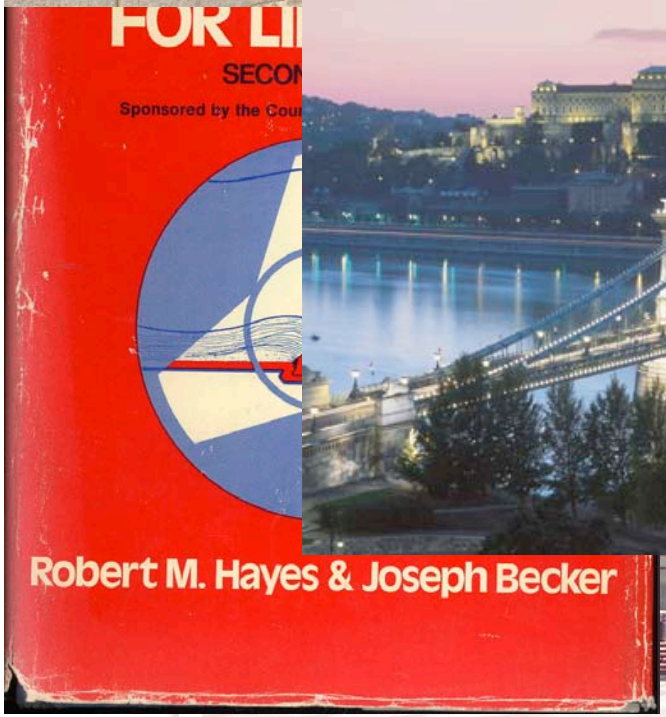
Everett Rogers

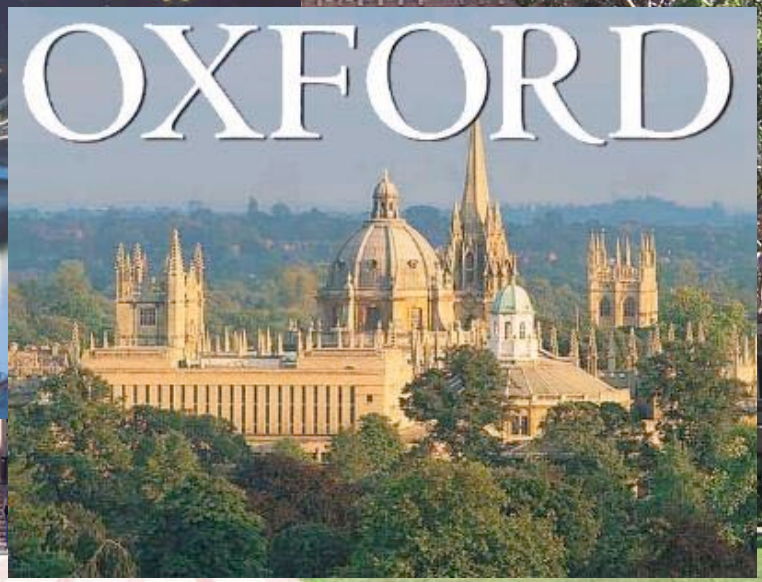
A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF COMMUNICATION
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Christine Louise Borgman
September, 1983

UCLA arrival: 1983

Robert Hayes





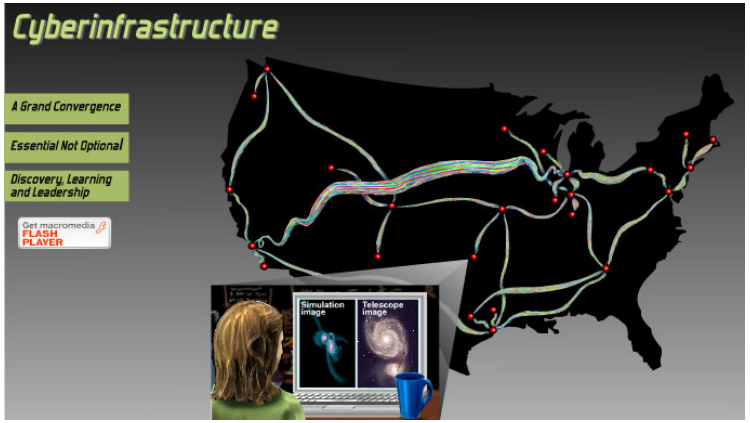
Information, Infrastructure, Internet

- Where are we now?

Kyoto, 2009

Cyberinfrastructure

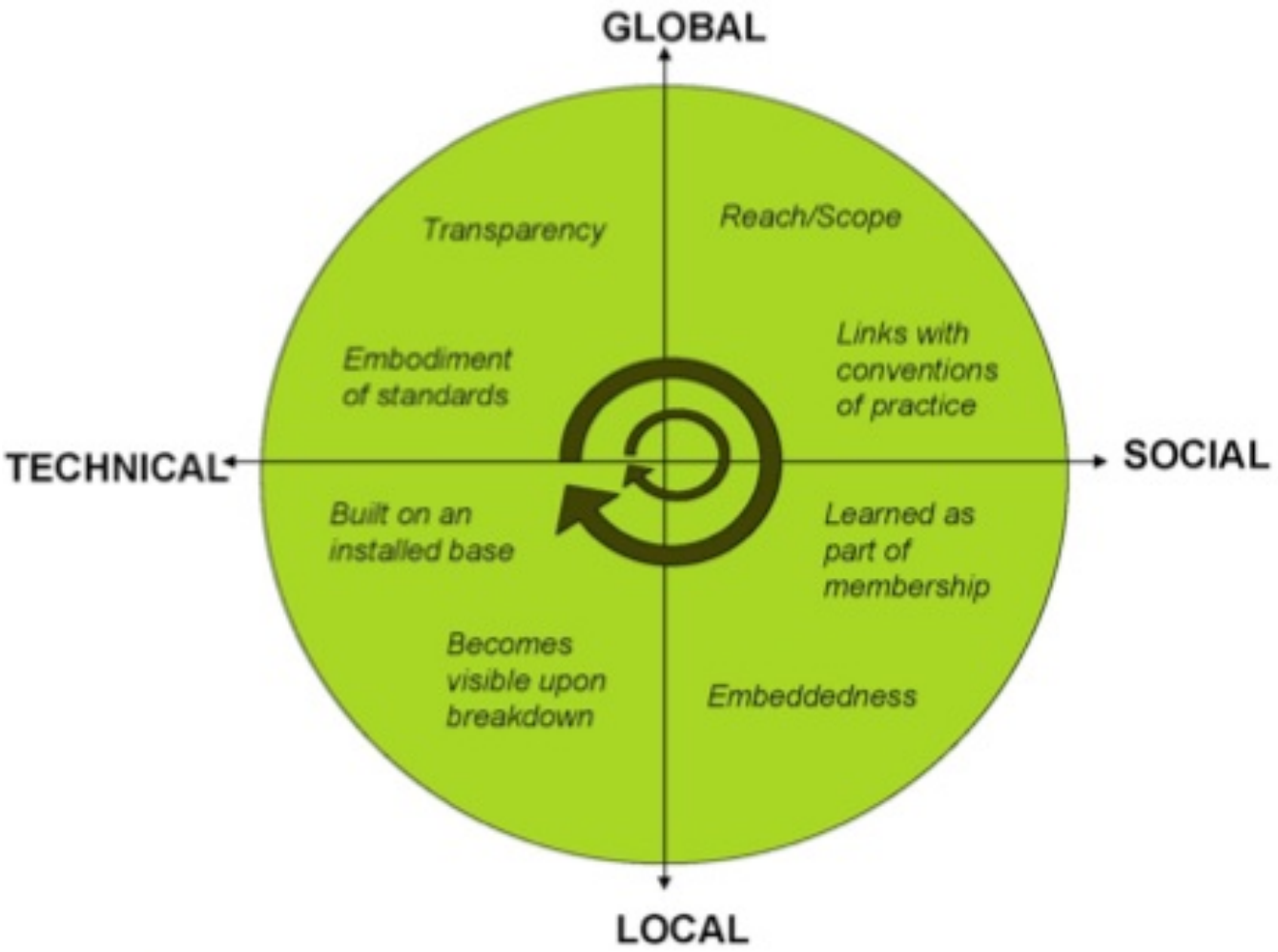
- Technology
- People
- Policy
- Scholarship
  - information-intensive
  - data-intensive
  - distributed
  - collaborative
  - multi-disciplinary



http://www.nsf.gov/news/special_reports/cyber/images/noflashintro.jpg

Tony Hey and Dan Atkins at Caltech, 2007

Infrastructure

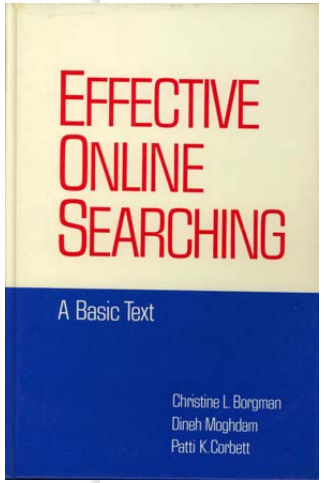


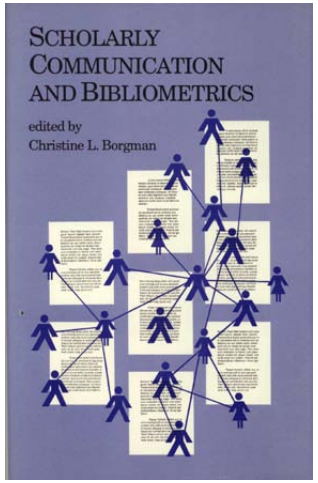
Star, S. L. & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. Information Systems Research, 7(1): 111-134.

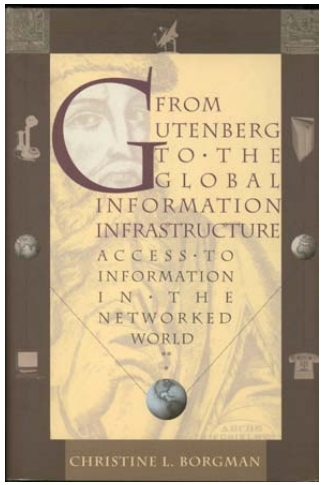
Figure by Florence Millerand, from: Edwards, P. N., Jackson, S. J., Bowker, G. C. & Knobel, C. P. (2007). Understanding Infrastructure: Dynamics, Tensions, and Design. National Science Foundation: University of Michigan. NSF Grant 0630263.http://hdl.handle.net/2027.42/49353

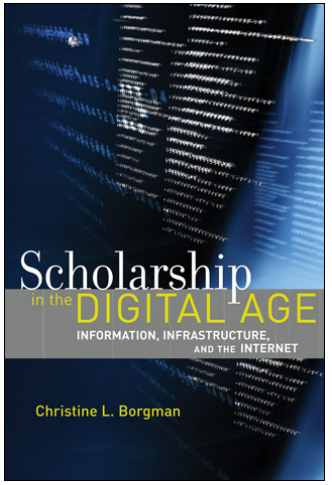
Infrastructure trends: 1970s-2010s

1. Closed to open
2. Static to dynamic
3. Readers to authors
4. Publications to data









1. Closed to open: 1970s-1990s

- Closed network
  - Research community
  - Bibliographic services
  - NREN
- Closed standards
  - Cataloging rules
  - MARC formats
  - Integrated library automation systems

http://americanhistory.si.edu/collections/comphist/objects/trs80l.jpg

1. Closed to open: 1990s-

- Open network
  - Academic and research communities
  - Commercial and public services
  - Commodity internet
- Open standards
  - Internet protocols
  - WC3 standards
  - Operating systems

http://www.gsd.harvard.edu/inside/building_services/Sustainability.html

1. Closed to open: Today

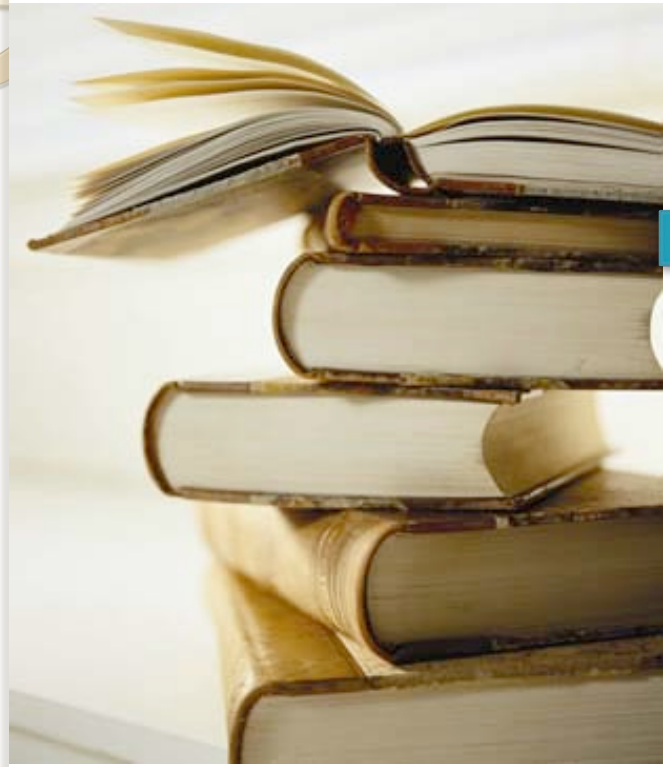
- Mixed open and closed network
  - Academic community
  - Public / private services
  - Commodity internet
  - App world

http://blogs.sun.com/dennisding/entry/sun_open_standards_definition_checklist

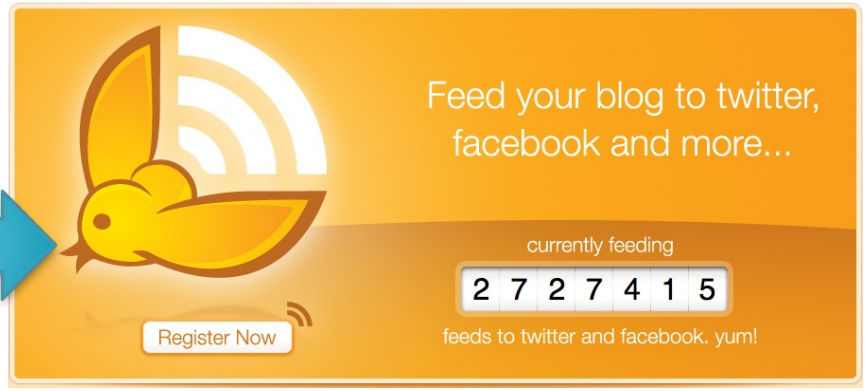
- Mixed open and closed standards
  - Internet protocols
  - WC3 standards
  - Operating systems
  - Application platforms

http://www.flickr.com/photos/gaetanlee/157465326/

2. Static to dynamic



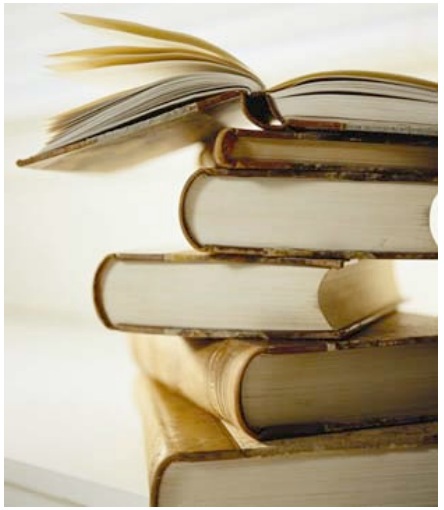
http://www.artsjournal.com/bookdaddy/2008/07/



http://twitterfeed.com/

2. Static to dynamic: 1970s-1990s

- Static content
  - Published documents
  - Databases grow incrementally
- Static context
  - Search results based on query
  - Same search yields same results
    - Anyone
    - Anywhere



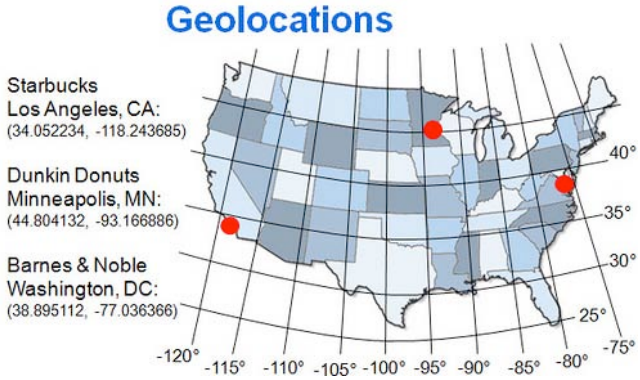
http://www.artsjournal.com/bookdaddy/2008/07/

2. Static to dynamic: 1990s-

- Dynamic content
  - Multiple versions of documents
  - Websites change constantly

http://www.flazh.de/en/404.htm

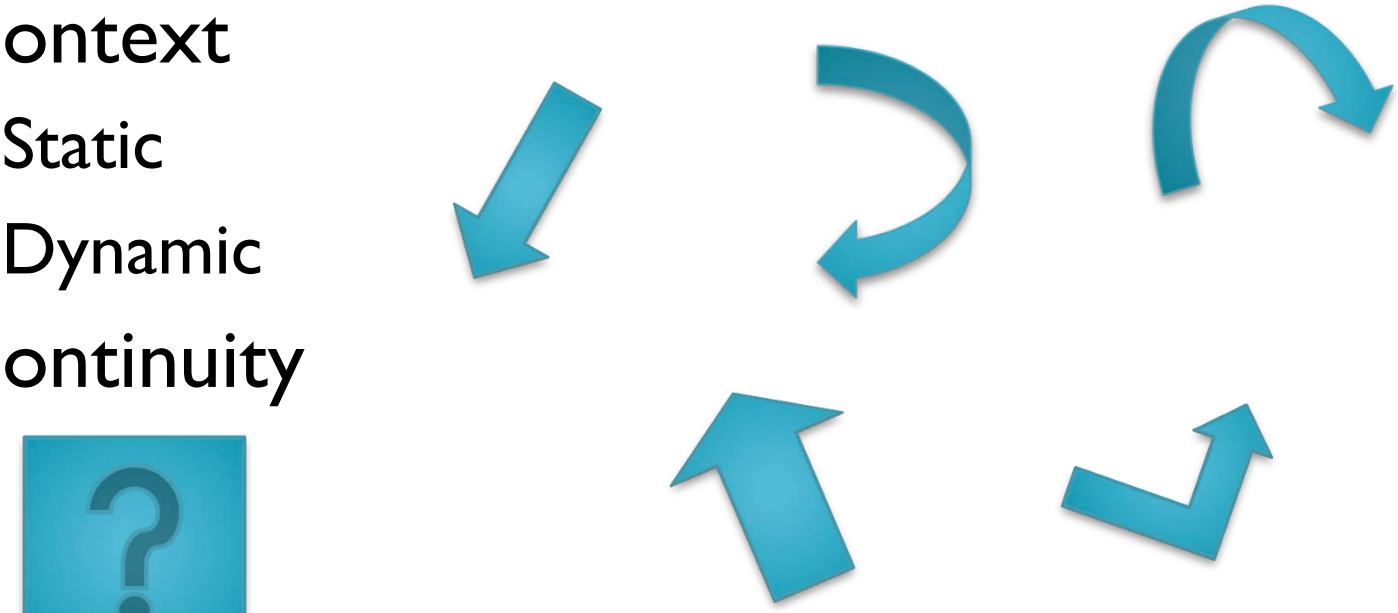
- Dynamic context
  - Search results vary by profile and location
  - Same search yields different results
    - Anyone
    - Anywhere



http://www.flickr.com/photos/silvery/5322139453/

2. Static to dynamic: Today

- Content
  - Static
  - Dynamic
- Context
  - Static
  - Dynamic
- Continuity



The Scholarly Journal, 1665-2011

PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD

Vol I.
For *Anno* 1665, and 1666.

In the *SAVOY*,
Printed by *T. N.* for *John Martyn* at the Bell, a little without *Temple-Bar*, and *James Allestry* in *Duck-Lane*,
Printers to the *Royal Society*.

3. Readers to authors: 1970s-2000s

- Reader services
  - Libraries
  - Finding information
  - Local clientele
- Author services
  - Publishers
  - Producing information
  - Global clientele

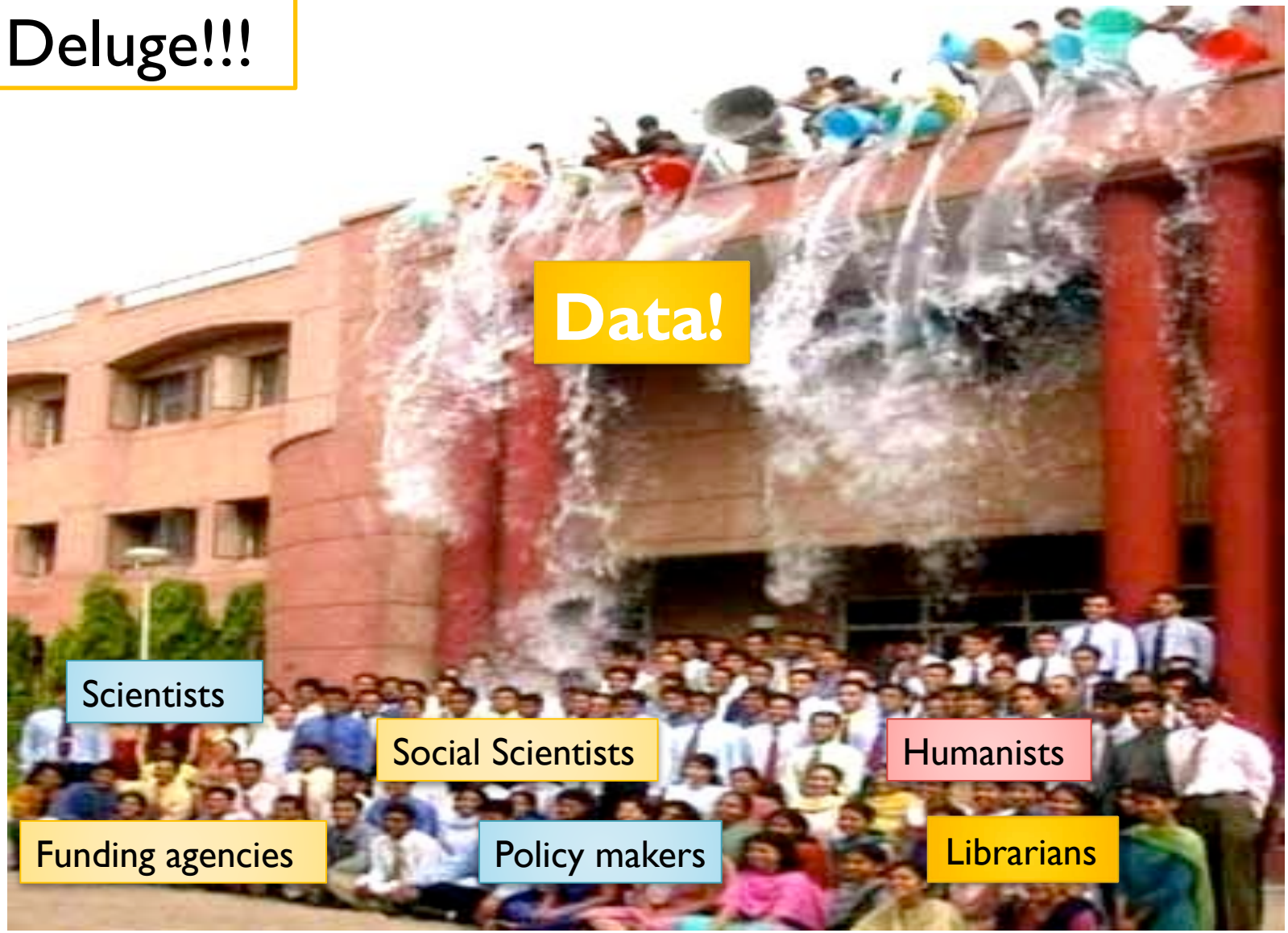
Ephesus, 2007

3. Readers to authors: Today

- Author DIY
  - Copy editing
  - Publication-ready formatting
  - Negotiating copyright
  - Supplemental materials
  - Institutional deposit
  - Maintaining access to data…
- Universal authorship
  - Email, blogs, tweets, Facebook…

http://smalllivingjournal.com/category/issue-4-do-it-yourself/

Deluge!!!

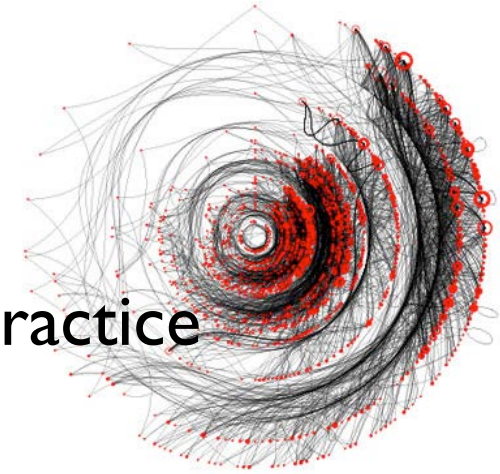


http://www.guzer.com/pictures/suprise_suprise.jpg

4. Publications to data: 1970s-1990s

- Publications
  - Types: Journals, books, papers…
  - Role: product at end of project
  - Social structure: Peer review, citation

- Data
  - Types: heterogeneous
  - Role: process
  - Social structure: embedded in practice



http://pingmag.jp/2007/03/23/infosthetics-form-follows-data/

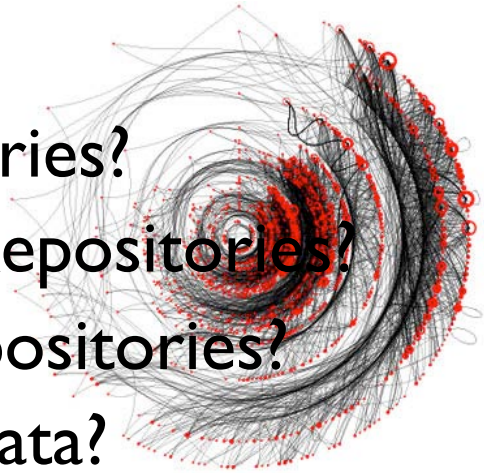
4. Publications to data: 2000s-

- Publications
  - Types: Journals, books, papers…
  - *Role: products throughout project*
  - Social structure: Peer review, citation
  - *Dissemination: libraries, blogs, tweets…*
- Data
  - Types: heterogeneous
  - *Role: process and product*
  - Social structure: embedded in practice
  - *Dissemination: publications, websites, repositories…*

http://pingmag.jp/2007/03/23/infosthetics-form-follows-data/

4. Publications to data: Today

- Publications
  - Peer review: Publishers
  - Cataloging: Libraries
  - Access: Libraries and search engines
  - Credit by citation: Authors
- Data
  - Peer review: Journals? Repositories?
  - Metadata: Authors? Libraries? Repositories?
  - Access: Authors? Libraries? Repositories?
  - Credit by citation: Reusers of data?



http://pingmag.jp/2007/03/23/infosthetics-form-follows-data/

Information, Infrastructure, Internet

- Where might we go from here?

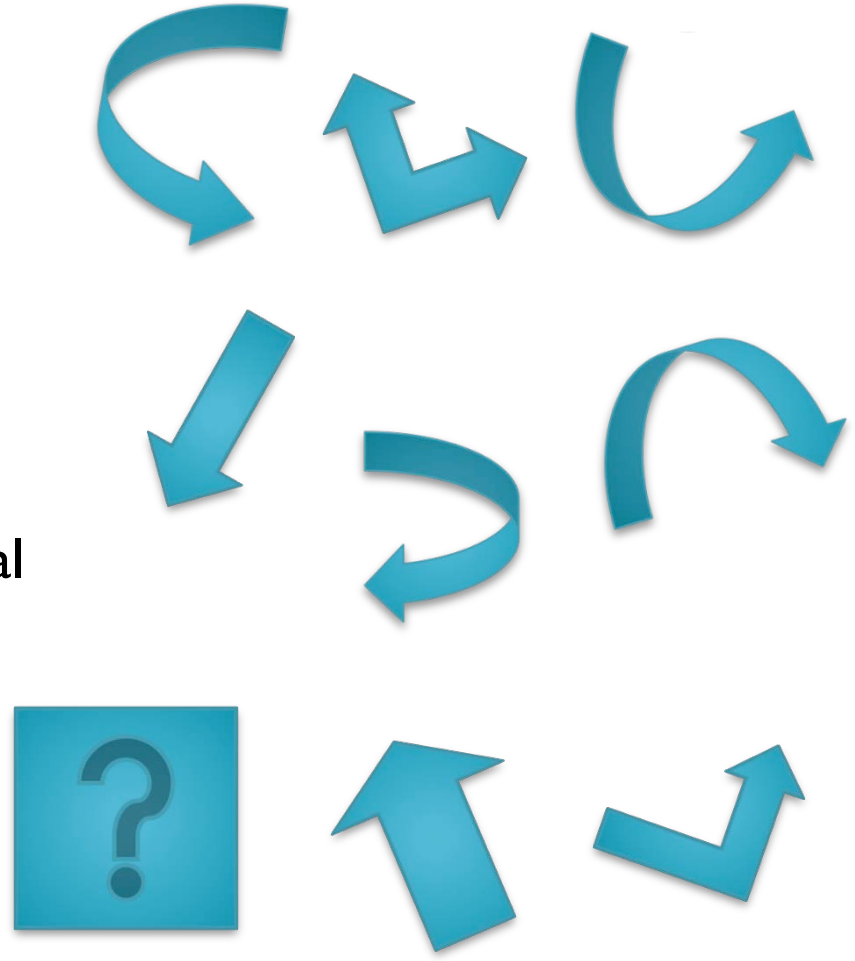
Grand Challenges for Information, Infrastructure, and the Internet

I. Take back information retrieval
II. Engage the information lifecycle
III. Distribute the architecture
IV. Match policy to incentives

Kanazawa Airport, Japan, 2009

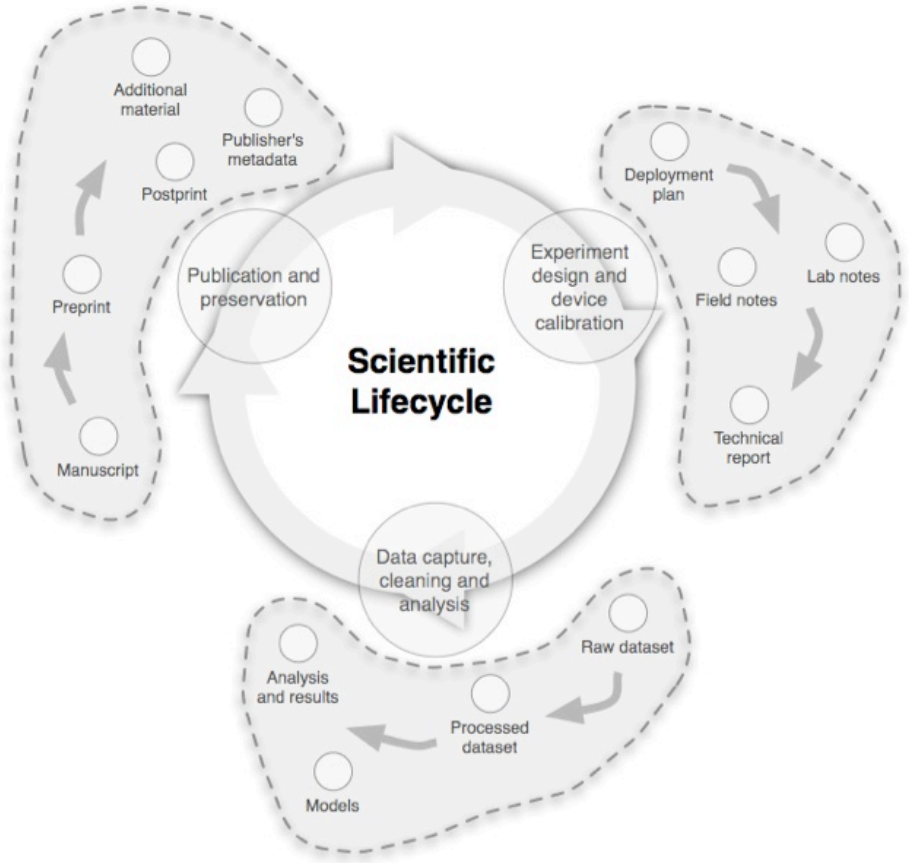
I. Take back information retrieval

- Today
  - Generic search engines
  - Bibliographic services
  - Discrete data collections
- Tomorrow
  - Discoverability
    - Generic search
    - Specialized search
  - Organization and retrieval
    - Aboutness
    - Linking related objects
  - Continuity
    - Reproducibility
    - Trust



II. Engage the lifecycle

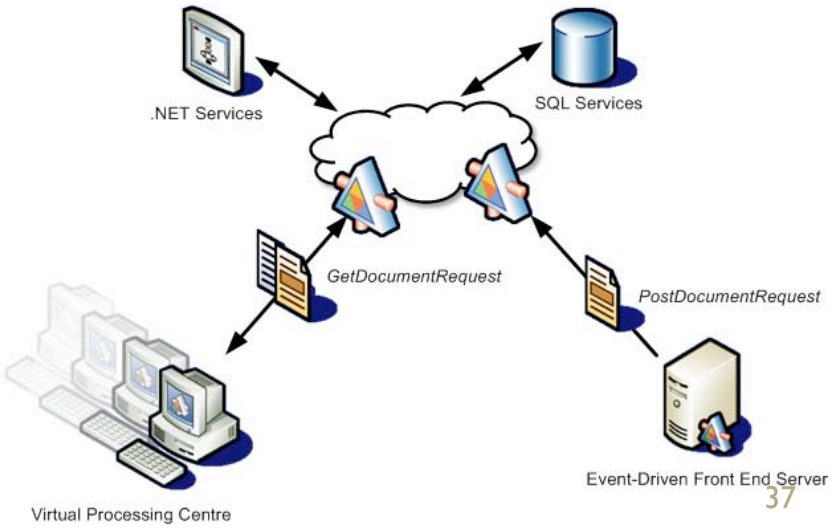
- Research design
- Data
- Metadata
- Writing
- Curation
- Use and reuse
- Reproducibility



Pepe, A., Mayernik, M. S., Borgman, C. L. & Van de Sompel, H. (2010). From Artifacts to Aggregations: Modeling Scientific Life Cycles on the Semantic Web. Journal of the American Society for Information Science and Technology, 61(3): 567–582.

III. Distribute the architecture

- Data scales fastest
- Select and filter
- Use surrogates for discovery
- Move computation to the data
- Share access *and* assets



http://geekswithblogs.net/EltonStoneman/archive/2008/12/02/cloud-distributed-computing.aspx

IV. Match policy to incentives

- Data curation: means or end?
  - Reuse
  - Reproducibility
- Data management is
  - Expensive
  - Poorly rewarded
  - Highly inconsistent
- Selection matters
- Stewardship matters

London, 2005

Borgman, C. L. (2010). Research Data: Who will share what, with whom, when, and why? China-North America Library Conference, Beijing. http://works.bepress.com/borgman/238.

Grand Challenges for Information, Infrastructure, and the Internet

I. Take back information retrieval
II. Engage the information lifecycle
III. Distribute the architecture
IV. Match policy to incentives

Kanazawa Airport, Japan, 2009

Thanks to co-authors and funders

Research Libraries Group, OCLC, Council on Library Resources, Sloan Foundation, National Science Foundation, UCLA, Microsoft Research