Never Optimize[1]: Building & Managing a Robust Cyberinfrastructure

Cal Lee, University of North Carolina at Chapel Hill
2006-09-11
Prepared for History and Theory of Infrastructure Workshop on 2006-09-28
Strategies Cluster - Organizational Capacities

The builder and keepers of a cyberinfrastructure (CI) must confront a major tension between two goals (1) an architecture to pushes the edge of the technical envelope of systems – e.g. optimally efficient processor use, storage use, input/output speeds -- and (2) the preservation of digital information that depends on those systems for trustworthy, meaningful and useful access over the long term.[2] The actors involved in the design and management of the components of cyberinfrastructure will not be well-served by fully optimizing to current conditions.

The NSF Cyberinfrastructure Council (NCC) offers the following principle: "Provide a framework that will sustain reliable, stable resources and enable the integration of new technologies and research developments with a minimum of disruption to users." A CI must "evolve" over time.[3] As opposed to systems, which are often designed to operate within a relatively bounded set of conditions, infrastructures have a scope that reaches beyond any single situation or context. Rather than being designed and implemented in one feel swoop, infrastructures are generally built as incremental advances on top of an installed base.[4] Rather than attempting to optimize for one specific context, infrastructure building will benefit from "robust action"[5] or "robust design,"[6] which is effective in the short-term but also sufficiently flexible to remain effective in a wide range of possible future contexts. Limiting the interdependencies between subsystems can also make a design more robust against disruptions from the environment[7].

One of the essential "reliable, stable resources" of a CI is "long-lived data,"[8] which requires "systematic archiving and curation."[9] The NCC points out that "research collections [originally developed to serve only short-term work group needs] may evolve over time to become resource and/or reference collections," which have longer periods of retention and thus require higher long-term stewardship commitments.[10] Such a transition is much more likely to be successful if considerations of interoperability and sustainability were built into the CI in which all collections are managed. It can be very difficult, if not impossible, to tack digital preservation considerations onto an existing system long after the information in the system was created.   Data archives should not be locked into one particular combination of hardware and software, but should instead make

---

[1] I borrow this phrase from Clay Shirky, who presented it during a meeting of the National Digital Information Infrastructure and Preservation Program (NDIIPP) in April 2003.

[2] Within a digital curation context, "long-term" is "A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository," which can often be as little as 4 or 5 years. *Reference Model for an Open Archival Information System*, Consultative Committee for Space Data Systems, 2002.

[3] "NSF's Cyberinfrastructure Vision for 21st Century Discovery," National Science Foundation, January 20, 2006, p.6, 9.

[4] David Alexander, "Infrastructure Evolution and the Global Electronic Marketplace: A European It User's Perspective," in *Standards, Innovation and Competitiveness: The Politics and Economics of Standards in Natural and Technical Environments*, Richard Hawkins, et al, ed (Brookfield, VT: Edward Elgar, 1995): 86-92; Susan Leigh Star and Karen Ruhleder, "Steps toward an Ecology of Infrastructure: Design and Access for Large Information Spaces," *Information Systems Research* 7, no. 1 (1996): 111-34.

[5] Eric Matheson Leifer. Actors as Observers: A Theory of Skill in Social Relationships (New York: Garland, 1991); John F. Padgett and Christopher K. Ansell, "Robust Action and the Rise of the Medici, 1400-1434," *American Journal of Sociology* 98, no. 6 (1993): 1259-319.

[6] Andrew B. Hargadon and Yellowlees Douglas, "When Innovations Meet Institutions: Edison and the Design of the Electric Light." *Administrative Science Quarterly* 46, no. 3 (2001): 476-501.

[7] Herbert A. Simon, "The Architecture of Complexity," *Proceedings of the American Philosophical Society* 106 (1962): 467-82.

[8] NSF Cyberinfrastructure Council, p.6, 8.

[9] Daniel E. Atkins, et al, "Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure," National Science Foundation, 2003, p.11

[10] NSF Cyberinfrastructure Council, p.17.

extensive use of redundancy[11]; diversity in both technological approaches[12] and business models[13]; abstraction; virtualization[14]; modularity[15]; detailed descriptive and administrative metadata beyond that which is required for immediate use; and the development and adoption of open standards in way that is attentive to the need for flexibility[16]. All of the above approaches run counter to those best suited to the purpose of "squeezing every last bit or cycle" out of current hardware and software to address a task that's current at hand. Meaning is expensive; transferring meaningful information across social or technological contexts requires additional resources. One of the fundamental challenges is that developers tend to focus on the task at hand (and possibly its fit within an architecture in place at the work group or organizational level), rather than seeing themselves as builders of an "infrastructure."[17]

Long-term curation of data within a CI requires not only robust artifacts and computer systems but also social systems that can both withstand and benefit from changes in the environment. The professions and organizations involved in data curation should also be cautious not to fall into a competency trap[18] of only being able to solve yesterday's problems. Professional abstractions[19] should be specific enough to solve specific problems but also robust to changes in the environment. Actors in this space should strive for "requisite variety"[20] in their repertoire of capabilities and "absorptive capacity,"[21] and they should actively monitor the environment for changes to both the ICT landscape and stakeholder needs/expectations. History suggests that the institutions responsible for information curation that are able to persist over long stretches of time are those that are able and willing to adjust their practices to fit changing funding models and use scenarios.[22] In short, "Long-term digital archiving requires systems, institutions, and business models that are robust enough to withstand technological failures, changes in institutional missions, and interruptions in management and funding."[23] Curation of data within a CI should ensure the integrity and fixity (or consistent reproducibility) of the information[24], while also supporting the technical malleability and "interpretive flexibility"[25] of the systems upon which the information resides. The data archives that support a robust CI must be "locked in" to a commitment to preserve collections of data, while not being locked in – or ever fully optimized – to one combination of devices or approaches for preserving it.[26]

---

[11] Petros Maniatis, et al, "The LOCKSS peer-to-peer digital preservation system," *ACM Transactions on Computer Systems* 23, no. 1 (2005): 2-50.

[12] David S.H. Rosenthal, et al, "Requirements for Digital Preservation Systems: A Bottom-up Approach," *D-Lib Magazine* 11, no. 11 (2005).

[13] NSF Cyberinfrastructure Council, p.19.

[14] Richard Marciano and Reagan Moore, "Technologies for Preservation," in *Managing Electronic Records*, Julie McLeod and Catherine Hare, ed (Facet Publishing 2005), 81-100.

[15] Carliss Y. Baldwin and Kim B. Clark. *Design Rules. Vol. 1: Power of Modularity* (Cambridge, MA: MIT Press, 2000); Richard N. Langlois and Paul L. Robertson. "Networks and Innovation in a Modular System: Lessons from the Microcomputer and Stereo Component Industries," *Research Policy* 21, no.4 (1992): 297-313.

[16] Ole Hanseth, et al, "Developing Information Infrastructure Standards: The Tension between Standardisation and Flexibility," *Science, Technology and Human Values* 21, no. 4 (1996): 407-26; Eric Monteiro, "Scaling Information Infrastructure: The Case of Next Generation IP in Internet," *The Information Society* 14, no. 3 (1998): 229-45; Tineke Egyedi, "Infrastructure Flexibility Created by Standardized Gateways: The Cases of XML and the ISO Container," *Knowledge,Technology & Policy* 14, no. 3 (2001): 41-54.

[17] Paul N. Edwards, "Y2K: Millennial Reflections on Computers as Infrastructure," *History and Technology* 15 (1998): 7-29.

[18] Barbara Levitt and James G. March, "Organizational Learning," *Annual Review of Sociology* 14 (1988): 319-40.

[19] Andrew Abbott, *The System of Professions: An Essay on the Division of Expert Labor* (Chicago, IL: University of Chicago Press, 1988).

[20] Karl E. Weick, *Sensemaking in Organizations* (Thousand Oaks, CA: SAGE Publications, 1995).

[21] Wesley M. Cohen and Daniel A. Levinthal, "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly* 35, no. 1 (1990): 128-52.

[22] Margaret Hedstrom and John Leslie King, "Epistemic Infrastructure in the Rise of the Knowledge Economy," In *Advancing Knowledge and the Knowledge Economy*, Brian Kahin and Dominique Foray, ed (Cambridge, MA: MIT Press, 2006).

[23] Margaret Hedstrom et al, "It's About Time: Research Challenges in Digital Archiving and Long-Term Preservation: Report on a Workshop on Research Challenges in Digital Archiving: Towards a National Infrastructure for Long-Term Preservation of Digital Information," National Science Foundation and Library of Congress, 2003, p.vii.

[24] David M. Levy, Scrolling Forward: Making Sense of Documents in the Digital Age (New York: Arcade, 2001).

[25] Trevor J. Pinch and Wiebe E. Bijker, "The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other," *Social Studies of Science* 14, no. 3 (1984): 399-441.

[26] Ole Hanseth and Kalle Lyytinen, "Theorizing About the Design of Information Infrastructures: Design Kernel Theories and Principles," unpublished manuscript, 2005.