

Concluding Remarks and Observations

“May all your problems be technical” [135]

Jim Gray, 1998 ACM Turing Award Winner

When the NSF, DARPA, and other US and international funding agencies began large-scale funding of digital library research in the early 1990’s, they were motivated by a number of goals. As I have described, the work on the first goal, stimulating basic research in networking, security, databases, information retrieval, and other areas, was quite successful. These largely *technical* endeavors produced a number of results that rapidly evolved from initial research prototypes to technology that was deployed and used on a global scale. Work on another largely technical goal (which arose largely in the context of DLI-2), the deployment online of a number of culturally, intellectually, and historically important digital collections, was also quite successful.

However, another highly-promoted goal, the development and large-scale deployment of new network-based information infrastructure, met with a number of obstacles, and the results were significantly short of success. The obstacles to success were not technical. In fact, many of the proposed solutions were technically sound and provided rich functionality (often greater than that which we see on the web). The actual barrier to success was the reality that infrastructure, and most notably information infrastructure, is a deeply *sociotechnical* phenomenon. As scholarship in

Science, Technology, and Society (STS) and numerous infrastructure studies have shown, the development and spread of infrastructure entails complex interactions among the technology, the people and organizations who will use and be impacted by the technology, the social norms that govern the interaction of people, organizations, and technical artifacts, and co-existing technologies.

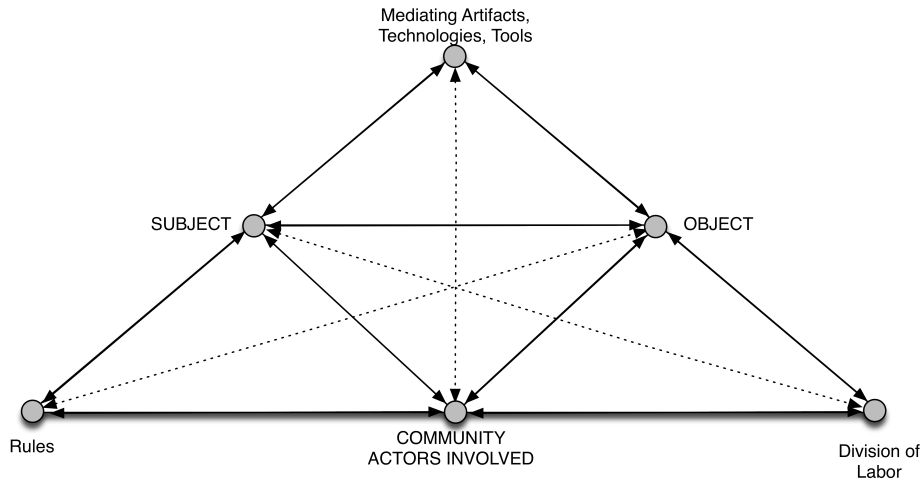


Figure 64 - Activity system

This complexity is the subject of the STS theories and frameworks described in Chapter 4. The notion of an activity system, illustrated in that chapter and repeated here in Figure 64, shows the mediated nature of technical change, especially one as broad reaching as information infrastructure. Activity Theory describes how technology is just one mediating factor in activity transactions and how a technology that does not mesh well with other mediating factors will be *disrupted*; modified or discarded. This is especially true for infrastructure technology, which should be non-intrusive and effectively invisible.

Because of a variety of factors described in Chapter 2, most digital library research effectively ignored this reality. Focusing mainly on technical goals, these researchers

assumed that existing, well-established norms of institutionally-based information organization and management would persist as the context for deployment of new search engines, scanning and display technologies, rights management mechanisms, and other technical advances. The digital library community was, of course, not alone in this myopia. Other examples include the entertainment industry, newspaper, publishers, and even the computer industry (e.g., Microsoft), all of which expected that digital information technology would somehow seamlessly mesh into their existing ways of doing business.

In reality, the act of putting information online and giving people almost universal access to that information dramatically disrupts every part of the activity network for virtually all information-related activities (i.e. every thing) that people do. As Yochai Benkler [48], Lawrence Lessig [335] and others have noted, it has a dramatic “democratizing effect”. The history of the web and the increasingly profound changes in Web 2.0 demonstrate the impact of this democratization. Based on a few relatively simple protocols and standards and constrained by virtually no rules, the web we use today is essentially an organic development – a creation from within the web itself, rather than defined by an institution, standards board, or funding agency. In the terminology of Activity Theory and Actor-Network Theory it represents a self-stabilization (and perhaps self-optimization) of the activity network in response to the multiple interactions of networked information technology with other components of the network.

Notably, these profound changes were enabled rather than determined by the web technology invented by Tim Berners-Lee in 1989. Those simple architectural components - resources, URIs, HTML, and HTTP - provided the basis for the rather simple Web 1.0 "document Web and, relatively unchanged, were later used as the

basis of the Web 2.0 "social web" that has so transformed the way we live, learn, communicate, and consume. Indeed, this basic web technology has not determined the nature of the information environments built upon it. Instead, the multiple applications and evolving forms demonstrate both the flexibility of the underlying technology and the shaping influences of different social contexts. We can expect in the future that these shaping factors will produce further unexpected changes in the way we interact with information.

In some ideal world, an analysis such as this would lead to a set of prescriptions for future efforts. In reality, though, the self-stabilization process describe above is probably non-deterministic and not attainable by tweaking some variables.

Nevertheless, there is certainly room for future research on heuristics on how to improve our approaches to infrastructure. Hopefully, the Data Conservancy project I describe in Chapter 13 will provide some progress towards developing these heuristics. In the meantime, we can expect to observe a continuation of the unpredictable and dramatic transformations in technology and information that we have experienced over the past two decades. As we work to find our way through these changes and choose alternatives, we may have to rely on the words of the enigmatic Yogi Berra who said "*when you get to a fork in the road, take it.*"¹⁵⁶.

¹⁵⁶ http://en.wikiquote.org/wiki/Yogi_Berra