Comparative Interoperability Project:
Collaborative Science, Interoperability Strategies, and Distributing Cognition

Florence Millerand[1], David Ribes[2], Karen S. Baker[3], and Geoffrey C. Bowker[4]

[1]LCHC/Science Studies
University of California,
San Diego
La Jolla, CA 92093, USA
1.858.534.6828
fmillera@ucsd.edu

[2]Sociology/Science Studies
University of California,
San Diego
La Jolla, CA 92093, USA
1.858.534.4627
dribes@ucsd.edu

[3]Scripps Institution of
Oceanography
University of California,
San Diego
La Jolla, CA 92093 USA
1.858.534.2350
kbaker@ucsd.edu

[4]Center for Science,
Technology & Society
Santa Clara University
Santa Clara, CA 95053
USA
1.408.551.6058
gbowker@scu.edu

This poster will present the Comparative Interoperability Project, a recently initiated study that provides a situated social and organizational comparison of three scientific information infrastructures deploying different approaches to data interoperability. Data interoperability refers to achieving data assemblage and sharing in order to enable reuse by various communities of users, across diverse disciplines and across extended periods of time.

We analyze the practical activity of deploying interoperability approaches through three case studies: GEON, a cyberinfrastructure project for the U.S. geo-sciences (Keller, 2003); the Long-Term Ecological Research Program, a federated information infrastructure directed toward ecological sciences (Hobbie et al., 2003); and Ocean Informatics, a nascent environment for the ocean sciences based at Scripps Institution of Oceanography (Baker et al., 2005).

While the issue of interoperability has remained primarily addressed as a technical one – e.g. in terms of a strategic choice of technical standard for data coding (e.g. classifications, metadata, ontology), we consider interoperability strategies as specific configurations of technical commitment, community involvement, and organizational structure. In other words, while deploying a specific interoperability strategy in terms of technical direction, an information infrastructure project also unfolds strategies of community mobilization and organizational arrangement (Kling and Scacchi, 1982).

Interoperability strategies develop in a large set of activities including planning, negotiation, and decision-making processes among heterogeneous actors and organizations within which various artifacts are mobilized, created, modified, and so forth. Typically, one could analyze the (passive) configurations of distributed cognition which are supported or not by the new technologies (Heath and Luff, 1992; Hughes et al., 1994). We analyze the emergent infrastructures rather as an active *process of cognition distribution*. We choose this second lens to analyze the development of interoperability strategies in the three scientific information infrastructure projects that we study. Following the scientists, technologists, data and information managers at work, we observe how, in their attempts to enact interoperability, they are actively engaged in a

process of distributing cognition between individuals, groups of individuals and artifacts. This dynamic analysis in terms of process of cognition distribution helps to understand the specificity of the configurations observed in terms of simultaneous definition of technological solutions, creation or modification of organizational structures, and involvement of communities.

Three levels of cognition distribution

In recent years, the analysis of interaction and coordinated work situations has enjoyed a renewal of its analytic categories thanks to the consideration of the ecological dimension of action: now understood inside distributed environments, populated with artifacts and technical mediations.

Research on situated action and distributed cognition have focused on the interactions between individuals on one hand, but also on the relations between individuals and collectives to their environment on the other hand, to propose new accounts of organized action (Hutchins, 1995; Suchman, 1987).

Science and technology studies have contributed to this renewal in analytic categories through laboratory studies and analysis of scientific work in complex institutional settings. Laboratory studies have revealed the distribution of cognitive activity within social networks, individuals and 'inscriptions' as a large part of scientific activity (Latour, 1987). Studies of cooperative processes in scientific work have disclosed the creation of artifacts or 'boundary-objects' in a collective course of action as methods of problem solving in heterogeneous and distributed environments (Star and Grisemer, 1989; Star, 1989).

In addition, some works have offered new ways to account for the role of material and symbolic entities in human action and behavior, arguing to include them fully in the analysis of human activities (Latour, 1983, 1992). In this perspective, objects – or artifacts – 'prescribe' a range of user's behaviors through the incorporation of 'scripts' or action programs in their mechanism (Akrich, 1992). These prescriptions relate back not only to technical functions to make use of the artifacts, but also to values, duties, and even ethics, so that the morality – as well as the competence – is distributed between humans and non-humans.

In our analysis of the differing interoperability strategies deployed in cyberinfrastructure projects, we try to describe how such a distribution of moral qualities is also a distribution of cognitive qualities.

We distinguish three foci of attention in an attempt to account for the distribution of cognition that occurs in the context of deploying interoperability strategies:
  i) The distribution of cognition between members of a collective or between collectives;
  ii) The distribution of cognition between individuals and a material environment – the environment acting as a support of collective coordination;
  iii) The distribution of cognition through time (past and future).

These categories are not mutually exclusive, but rather serve as a heuristic for organizing our analysis. We present two examples of interoperability strategy

deployment drawn from the project case studies, and the related processes of cognition distribution that are involved.

Ontologies building as cognitive delegations

As a cyberinfrastructure project for the US geo-sciences, GEON aims at providing scientific data and resource sharing services to a broad range of disciplines to ensure a more integrated picture of earth processes. One challenge that this project faces is its capacity to integrate heterogeneous resources (e.g. datasets) and to provide flexible views on these resources so that different user communities can use them. The creation of ontologies has been chosen as a strategy for data interoperability. Briefly, ontologies are formal conceptual maps of domain knowledge that, when linked to datasets, allow the user to navigate more easily in unfamiliar domain knowledge databases.

The activity of ontology building involves cognition distribution processes that occur at different moments across time. First, ontology production is situated within specific organizational settings that bring together geo-scientists experts and information technology specialists in order to achieve knowledge elicitation and codification. So this 'sample' of geo-scientists coupled with information technologists is responsible to achieve the work of drawing equivalences between disciplines on behalf of the entire community of geo-sciences. Second, the specified knowledge became encapsulated in ontologies. Thus, ontology holds the cognitive work of conceptual equivalence drawing and knowledge mediation while assisting the user to navigate in the database. In other words, the cognitive work of knowledge codification and elicitation has been delegated to a group of representatives, and the work of knowledge mediation to an artifact.

In short, this interoperability strategy deployment has involved processes of cognition distribution that become stabilized in organizational structures (workshops within which a sample of geo-scientists together with information technologists are in charge of creating knowledge models) as well as in technical interoperability solutions (ontologies that hold knowledge mediation).

Organizational arrangements as cognition distribution between collectives

Ocean Informatics is a nascent information infrastructure for the ocean sciences within a single oceanographic institutional department that aims at providing a set of resources including shared scientific data and a design studio for learning, tool sharing and participatory design. Within the deployment of its interoperability strategy, Ocean informatics is currently at the stage of community building.

Creating a sustainable environment for a community to explore and carry out the work of informatics constitutes the starting point of Ocean Informatics interoperability strategy. In addition, information managers work closely with ocean scientists to make informed decisions about what kinds of organizational work are coupled to technical choice, thus opening the possibility for reflexive community participation in their own transformation prior to technical implementation.

The design of organizational structures and work processes within the information managers' community has revealed a tension between two needs: the development of a

common conceptual framework and the development of technical solutions. In recognition of the differing topics and timeframes, both a reading group and a series of design groups have been created as distinct entities. The former presents a space dedicated to conceptual elaboration and discussions; and the latter provides places for the pragmatic work of technical development.

This organizational arrangement has involved the distribution of cognition between collectives of individuals, leading to the creation of two environments with differing cognitive qualities that have been embedded into specific organizational structures: a reading group for conceptual elaboration and designs groups for technical development.

Conclusion

While deploying interoperability strategies - as technical direction, community mobilization, and organizational structure, scientists, technologists, data and information managers appear actively engaged in processes of cognition distribution, so that a major part of their work consists precisely in the stabilization of these distributions, embedding them into technical artifacts, organizational arrangements and community representatives.

An understanding of these cognition distributions as processes rather than as given configurations allows for their analysis in a dynamic way as well as the analysis of the negotiations that have shaped them.

References

Akrich, M. (1992). The De-scription of Technical Objects. In W. E. Bijker & J. Law (Eds.), *Shaping Technology, Building Society. Studies in Sociotechnical Change* (pp. 259-263). Cambridge, MA: MIT Press.

Baker, K. S, Jackson, S. J., & Wanetick, J. R. (2005). Strategies Supporting Heterogeneous Data and Interdisciplinary Collaboration: Towards an Ocean Informatics Environment. Proceedings of the Hawaii International Conference for System Science, HICSS38.

Heath, C., & Luff, P. (1992). Collaboration and Control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms. *Computer Supported Cooperative Work* (1), 69-94.

Hobbie, J. E., Carpenter, S. R., Grimm, N. B., Gosz, J. R. & Seastedt, T. R. (2003). The US Long Term Ecological Research Program. *BioScience*, 53(2), 21-32.

Hughes, J., King, V., Rodden, T., & Andersen, H. (1994). *Moving out from the control room: ethnography in system design*. In Proceedings of CSCW'94 (pp. 429-439), Chapel Hill: ACM Press.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Keller, R. B. (2003). GEON (GEOscience Network)-A First Step In Creating Cyberinfrastructure for The Geosciences. *Electronic Seismologist*, July/August.

Kling, R., & Scacchi, W. (1982). The Web of Computing: Computer Technology as Social Organization. *Advances in Computer* (21), 1-89.

Latour, B. (1992). Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In W. E. Bijker & J. Law (Eds.), *Shaping Technology, Building Society: Studies in Sociotechnical Change* (pp. 225-258). Cambridge, MA: MIT Press.

Latour, B. (1987). Science in Action : *How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

Latour, B. (1983). Give Me a Laboratory and I will Raise the World. In K. Knorr-Cetina & M. Mulkay (Eds.), *Science Observed: Perspectives on the Social Study of Science* (pp. 141-170). London: Sage.

Star, S. L. (1989). The Structure of Ill-Structured Solutions: Heterogeneous Problem-Solving, Boundary Objects and Distributed Artificial Intelligence. In M. Huhns & L. Gasser (Eds.), *Distributed Artificial Intelligence* (pp. 37-54). Menlo Park: Morgan Kauffmann.

Star, S. L., & Griesemer, J. (1989). Institutional Ecology, 'Translations,' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939. *Social Studies of Science* (19), 387-420.

Suchman, L. (1987). *Plans and Situated Action. The Problem of Human-Machine Communication*. Cambridge: Cambridge University Press.