# Origins of Multi-Sector Scientific Collaboration: A Report on Research in Progress

Judith Weedman
University of California, Berkeley

paper presented February 1992
annual conference of the International Network for Social Network Analysis

**Abstract**

Sequoia 2000 is a research initiative funded by the Digital Equipment Corporation to develop large capacity object servers to support global change research. Existing hardware, software, network technology, and visualization techniques are inadequate to the task of handling the terabytes of data which global change researchers need to access and manipulate. The purpose of Sequoia 2000 is to develop the needed technology and to create an electronic repository in which researchers' data sets, programs, documents, and simulation outputs can be stored and made available to multiple users. Sequoia 2000 is a multidisciplinary, multi-campus, multi-agency project; researchers are from the fields of computer science, information science, and global change, and are located in private industry, universities, and state and federal agencies.

This is a report of the initial stage of phase one of a study exploring collaboration and technology transfer resulting from this large research initiative. This phase will examine motivations for collaboration at different stages of the project and by players with different relationships to it. Investigators' initial motivations for participation, collaborative behavior attributable to the project, and the extension of Sequoia technology to colleagues of the initial vinvestigators will be examined. This first report focuses on the inter-relationships among the goals of the various initial participants and the activity which brought the project into existence, and describes plans for the remainder of the project.

# 1 Introduction

Collaborative work is a preoccupation within both industry and academia. Collaborative work is thought to increase productivity and lead to greater innovation than a single researcher can achieve alone. University rhetoric supports interdisciplinary research and urges small campus departments particularly to collaborate with other units on campus. Industry promotes collaboration both at the level of engineering teams and at the level of multi-company consortia.

There have been a few suggestions that collaborative work may not be directly tied to increases in productivity or competitiveness. Research by Kraut, Egido, and Galegher on articles written in the discipline of psychology indicates that faculty who wrote primarily collaborative articles were no more productive than those who wrote primarily non-collaborative articles. Also, the faculty members tended to rate their individual articles higher than their co-authored ones; they felt the writing was clearer, more ingenious, of better quality, and also that the work itself was more central to the field and of higher theoretical and empirical quality. Nor did more objective measures show collaborative articles to be superior; both kinds of articles were cited an average of 6 times within 3 years of publication, and a panel looking at the writing specifically did not find either collaborative or solo articles to be better-written. Nor was time to publication decreased; both types of articles took 19-20 months from planning to writing a draft for publication.

Obviously, collaboration is essential in some fields, but this research in a field where both individual and collaborative work are possible indicates that there may not be as strong an intrinsic intellectual advantage as has been believed.

At the level of consortia of organizations, recent writers (Werner & Bremer, 1991) assert that neither our consortia nor those of foreign countries have paid off in spectacular ways. Many of the US consortia were formed in response to consortia in other countries, the Japanese Fifth Generation Computer Systems Project, for instance. One author asserts that Japanese research associations have tended to "pursue non-threatening and often low-priority products," and that while good research is being conducted, it is usually "incremental rather than radical, and remote from the commercialization process." (quoted by Werner and Bremer, 1991). In addition, many assert that US consortia have not contributed much to their member companies' competitiveness and have had very little impact. Others argue that it is too soon to make that evaluation.

A related problem in industry, both in-house and in multi-organizational collaboration, is that of technology transfer. The movement of information from research units to production units, whether the research unit is the company's own or a university or consortium collaboration, is not automatic or easy. Ideas and prototypes generated in one unit cannot simply be handed to another for implementation. Often what needs to be transferred is not a finished product but an idea-in- progress, by nature somewhat ambiguous and ill-defined (Irwin and More, 1991). The process involves communication of a particular way of defining the problem, of conceptualizing what could constitute a solution, methods of attack, acceptability of various kinds of compromises, and so on. Points of view and frames of reference have to be transferred as well as specific techniques.

The solutions proposed all involve improving communication across the boundary that separates R&D from production. Irwin and More suggest a model in which the boundary is conceptualized as a space rather than a line, an intermediate space which should be occupied by individuals who are neither fish nor fowl, but who instead have close ties to both R&D and production and spend time running between them literally and figuratively to build meanings and create a shared orientation.

Digital Equipment Corporation has chosen a strategy of requiring that every external research project they fund have an identified sponsor within the corporation, who will be that product's champion, thus facilitating the eventual transfer of the results of the project into the corporate environment.

Monsanto learned from its disappointing agreement with Harvard for joint biological and medical research that agreements by heads of organizations are not sufficient to make the transfer occur. While Monsanto provided a lot of money and a lot of good research was done at Harvard, "Monsanto was not able to pick up the results to its advantage" (McHenry, 1990). What's required, once the agreements are signed at the top levels, are daily contacts at the level of the working scientists and engineers. Technology transfer has been referred to as a contact sport; there has to be face-to-face interaction if the intangibles of innovation are to be communicated.

Information transfer is the primary issue in collaborative work and in technology transfer. Intra- and interorganizational boundaries separate not just administrative units, but clusters of goals, motivations, ways of defining problems, traditional methods of problem solution, and general cultures and mind-sets. Both the assets and the limitations of projects are likely to involve individuals' and organizations' abilities to transfer ideas and information across such boundaries.

## 2 Sequoia 2000

Sequoia 2000 is a research initiative which is rich in boundaries. It involves, at last count (February, 1992), 16 organizations distributed across four economic sectors – the primary funding corporation, Digital Equipment Corporation, plus three additional corporations, five University of California campuses, 2 other UC units (the San Diego Supercomputer Center and Scripps Institution of Oceanography), three federal agencies, and two state agencies.

The core of the project is the design of computer science solutions to data handling problems in global change research; the two Principal Investigators are a computer scientist from UC Berkeley and a global change scientist from UC Santa Barbara.

Earth science researchers have a massive data handling problem which is growing daily. Global change scientists accumulate and need to be able to analyze data in large quantities and of a variety of types. Typical data sets include topographic data at various levels of resolution, satellite scan data at various resolution levels, sensor data, and measurements of such things as snow grain size, chlorophyll concentration, depth of snow pack, vegetation, and more. Not only is there great variety in the kinds of data used, but the data are gathered over time, and in principle over all of the world's surface. The situation is more complicated yet because in addition to such observational data, researchers also use model-based simulation data, and the interaction between the two kinds of data is important.

The goal of Sequoia 2000 is to address these problems through advances in computer technology - hierarchical storage, networking, visualization techniques, and distributed database management. An electronic repository will archive the disparate data sets, routines for their analysis, and various documents such as technical reports.

# 3 Collaboration, technology transfer, and communication in Sequoia 2000

This study examines collaboration, technology transfer, and communication as they occur within the Sequoia project. This paper is an in-progress report on Phase 1 of the study, which is concerned with the origins of Sequoia and will gather baseline data for use in later phases. My report today is based on only a few interviews, so the caveat must be made that the data are incomplete and may even be found to contain some level of error as the study continues. In addition, changes in structure and commitments will take place as the project evolves.

The two central issues for this project are the goals and motivations of the organizations for participating in Sequoia, and the origins of the project. The method used was in-person interviews combining open-ended and standardized question formats.

## 3.1 Motivation and goals

**Digital Equipment Corporation.** Located on Massachusetts' high tech Route 128, Digital is a giant in the computer industry. A $14 billion company, it employs 120,000 people around the world. More than most large companies, Digital puts a priority on joint research projects with other organizations. Within their corporate research organization is a unit which establishes and carries out these projects, the External Research Program. At any given time, ERP has 100 or more projects going on world-wide, most of them with universities but also some with government labs and other research labs. Sequoia 2000 is by far the largest of Digital's external projects.

The primary goal was to acquire the knowledge and technology which would be developed in the course of the project. The External Research Program's purpose is to stimulate research and to participate in it, in order that DEC can benefit from the developments. However, for this project, Digital did not specify what area of computer technology should be addressed. A second goal was to create a project that would increase the company's understanding of collaboration and technology transfer processes; DEC, like virtually all corporations, struggles to accomplish these processes successfully internally, and also has clients who have problems in these areas. Therefore, rather than being structured around a specific market activity or market interest, the RFP was designed to focus attention on broader issues of process as well as the technological issues.

**The people with the problem.** One of the two Principal Investigators is from the UC Santa Barbara Center for Remote Sensing and Environmental Optics. There are an additional 12 earth science researchers serving as Faculty Investigators, three at CRSEO, four at UCLA, 4 at Scripps, and one at UC Irvine.

These researchers saw in the project an opportunity to gain equipment which would allow them to do their own research faster and to manipulate their data in better ways than had been possible so far. Global change research is highly collaborative by nature, and Sequoia presented an opportunity to do more and better collaboration.

**The people with the solution.** Ten faculty members in computer science and information studies. Six, including one of the Principal Investigators, are at UC Berkeley, two at UC San Diego, and two at the San Diego Supercomputer Center.

- **Computer Science.** The primary motivation was the opportunity the grant would provide to do new research, to advance the field. Berkeley's CS department has had a tradition of big data base projects going back at least 20 years, and at the time the DEC RFP was being circulated, they were looking for their next big project. The hardware and software people both saw tertiary memory, involving storage hierarchies, as the obvious next progression of their work. CS at this time was holding a series of seminars which brought in individuals who had potentially interesting problems which could give them an application for the research they wanted to do.
- **Information Studies.** As more information of all kinds becomes digitized, as a wider range of objects comes to be seen as appropriate for electronic storage, emphasis in the field of library and information studies is shifting from purely bibliographic information to storage, retrieval, and use of information of any kind, with little regard for its format. The Faculty Investigators from the UC Berkeley School of Library and Information Studies see Sequoia 2000 as the library of the future. Participation gives them an additional opportunity to work on large-scale retrieval problems in the context of a wellfunded project. It also provides access to a wide range of data types and an opportunity to collaborate with researchers in other disciplines.

**National agencies.** The National Oceanic and Atmospheric Administration, the National Center for Atmospheric Research and the University of Colorado where it is located, and the US Geological Survey also engage in global change research. They have data sets which will be archived and will have access to the data sets of the other participants and will have access to a broader array of global change data through the electronic repository.

**State agencies.** The California Air Resources Board has precipitation chemistry measurements which will be archived in the electronic repository. The California Department of Water Resources is interested in the work of researchers at UC Santa Barbara on runoff, and is interested in geographic information systems for cropping patterns. They also committed to preparing a promotional tape for the Sequoia project.

**Other corporate partners.** Exabyte and Xerox/PARC are donating hardware and software to the project. ESL Corp will have observers involved because of their interest in possible applications to reconnaissance technology.

Technology transfer, visibility from association with the project, and incorporation of their technologies into future systems are primary goals for the other industrial partners.

### 3.2 Origins

In 1990, Digital's large-scale and well-known Project Athena was coming to the end of its funding period, and it was time to decide whether DEC would fund another big project, or take the money and distribute it to a number of smaller projects. The decision was made that it was valuable to have a large project in place, one which was not just different in scale, but different in kind from the smaller projects.

It took approximately a year to work out what sort of project they wanted to do. As noted earlier, ERP decided against specifying the content of the project. One difference between Sequoia and Athena is that in the early 1980s, when Athena was being put together, it was very clear what the next area for advances in computing technology was. The move was clearly going to be from mainframe to distributed systems, and there was consensus in the computing world generally that this was the area where the next big advances were going to take place. In 1990, it wasn't clear that there was a single area that was going to take off. Digital therefore made the decision to leave the specific nature of the research open. They did, however, have certain components they wanted any project to include. The general goal was to take a specific, real-world problem that was at the far edge of what's currently possible in computing technology, and design a solution. That solution should then in principle be generalizable to other applications areas, and to commercial products.

The dimensions which were identified as being desirable in this next project were (1) that it be interdisciplinary and multidisciplinary, (2) that it be radically collaborative, involving a real-world application as well as a computer science technical problem; they wanted an application-pull rather than a technology-push, (3) that it be a multi-campus project, and they would have been pleased if it had turned out to be multi-national as well. (4) They wanted the project to address an area in computer science and computer technology that would be one of the next really important areas, where strides needed to be taken and where the important products would be developed 5 to 10 years down the road.

Digital sent out RFPs through e-mail and word of mouth to the universities with which it had an established relationship. At this stage they requested short think-pieces, not formal proposals. Key people at this stage were the sales personnel who worked with the various universities; they were important in providing information about what DEC was looking for, and also in stimulating faculty members to get together and think about what they might do.

DEC got 30-some responses to the first request for think-pieces, and sifted them down to the five that looked the most solid and promising. These five groups were then asked to write major proposals. Three of them were from universities in Europe, two in the United States. A panel composed of senior engineers, senior engineering managers, and senior marketing people reviewed the proposals, heard presentations by each of the 5 groups, and made the decision to award the grant.

What made Sequoia the obvious choice was that not only did it involve talented people working on an interesting project, but it had an internal mechanism which would keep them working together. Many of the proposals suggested projects which were basically independent interests running in parallel. The different pieces didn't really need each other, and there was no intrinsic need for them to cooperate. In the Sequoia project, on the other hand, the primary collaboration – between the computer people and the global change people – was essential. The technical people needed the problem, and the global change people needed the solution.

At the time the informal Request for Proposals went out, Berkeley's computer science department was ready for a new database project. They had commenced a seminar series which brought in individuals with interesting problems which might represent areas for research. One speaker was from the Center for Motion Picture Study of the Academy of Motion Picture Arts and Sciences; the archive has miles of nitrate-based film and wants to digitize it. There are two reasons; nitrate-based film is volatile, and digitization would provide increased access. Manual indexing for the film-based medium gave access through director, major actor, and one or two other features; a database of digitized material would provide access through many features, such as minor actors and location. Another speaker was from the Human Genome Project, which uses databases in matching gene pairs. Another idea which surfaced at this time was to scan the contents of Berkeley's rare book library, which would provide an electronic database of about the right size. This idea disappeared when it was realized that current scanning technology would require removing the spines of the books.

Some time during that same semester, the UC Office of the President was approached by Digital with the RFP requesting that they submit a short think-piece about what the next important area of advance in computing technology would be and how the UC system might design a project in that area. Digital received at least three proposals from investigators at UC, including the one which eventually evolved into Sequoia 2000.

Simultaneously, the Digital account manager who serves the UC system was talking to 10-12 faculty members whom she knew about this

RFP, stimulating them to think about what they might submit.

During this same time period, the chair of the CS division met a UC Santa Barbara global change researcher while they were both serving on a national research council.

At that point, the necessary links were in place.

The two men who would be the Principal Investigators and a staff programmer spent one morning together and hammered out the basic shape of the proposal. Additional information was provided by the UC account manager from Digital, who knew what Digital was looking for and what kinds of projects were most likely to be funded. A first draft of the proposal was sent out through e-mail to a large number of people, in effect a straw proposal for them to react to. As new ideas came in, successive drafts were sent around, people rewrote sections and added new ones, until it was in final form. Many of the people who contributed to the proposal had never met one another.

## 4 Discussion

**Goal congruence.** Even though Sequoia was designed so that its two poles – the computer scientists and the earth scientists – need each other, the overlap of their goals is not total. The CS people are concerned first with the design of a system which has the capacity to provide access to data and routines at a relatively high speed. The global change people are concerned with using a system to do their work. The global change researchers would, no doubt, be happiest with a finished, polished, consumer-oriented product, and that's not what Sequoia is designed to produce. The desired advances in computing technology can, for a researcher, be gained at a much earlier stage than that of a finished, marketable product. The ground between the two outcomes will have to be negotiated. Digital may consider a third outcome desirable, in relation to its need to commercialize products developed through research.

**Levels of investment.** Wigand (1990) found in a study of microelectronics firms that corporations most often initiated collaborative projects with universities when they had a specific problem with which they needed help or when a personnel shortage was handicapping work on a pressing research issue. As noted, Sequoia did not originate from a well-focused market need; its goals were intentionally more diffuse and long-term than is common in such projects. It may be more difficult for a corporation to maintain commitment to a project of this nature than to projects with direct paths to market. The investment of Faculty Investigators varies as well. For many of the computer science researchers this project is their work, a direct line of development from prior research. For the global change scientists, on the other hand, time spent on Sequoia is separate from the trajectory of their work, an additional responsibility with anticipated but uncertain long-term benefits. The investment of most of the governmental agencies appears to be the lowest of all. Auster (1990) argues that the higher the level of resource investment and the greater the resource dependence, the more likely organizations are to continue established relationships. Sequoia was designed to create as much inherent inter-dependence among the participants as possible, but such relationships are dynamic, and change configuration as economic and social pressures shift.

**Social structure.** Social networks which pre-existed the project contributed to its development, but those initial networks may have been more limited in the case of Sequoia than is common. Wigand (1990) found in his study of interorganizational collaborative projects that in more than 70% of the cases, the industrial and university researchers had known or known of their colleagues for 3-5 years before they worked together on a project. The two key groups, the computer scientists and the earth scientists, were virtually unknown to each other before the project. As noted earlier, many of the key contributors to the proposal had never met. The central role in bringing these strangers into collaboration with each other seems to have been played by DEC personnel, who brokered the new relationships. Similar results were found by Johnson and Tornatsky (1981) in their study of NSF funded cooperative programs; NSF staff were often the linchpins in the collaborative relationships.

New links between individuals and organizations are imposed by the project, and others will spring up spontaneously within the project. These will affect the course of the project and may extend beyond the project into long-term collaborative relationships.

**Communication issues.** These issues are seen by many of the Sequoia investigators as central to the success of the project, and a great deal of thought has gone into devising ways to establish and maintain group cohesiveness. Wigand (1989) found that the greatest cost to universities involved in collaborations with industry was in travel time. In a project of this size, travel was not an attractive option, and a variety of communication media have been employed in parallel to replace face-to-face communication.

The Berkeley seminar instituted in 1991 has been continued, changing its role to become the symbolic centerpiece of the communication process. Once the project was underway, each seminar was videotaped and tapes were distributed to all project sites. This provided a unidirectional flow of information, but did not permit timely feedback or questions; the next stage was to utilize speaker phones and conference calls, with accompanying graphics distributed via fax to all sites. Negotiations are now underway to obtain teleconferencing equipment for the seminars. The weekly seminar is the only regular communication activity that involves all participants, and it serves an important function in group cohesion. Researchers present summaries or tutorials for those outside their disciplines and report on research in progress. People not connected with Sequoia who have related interests are brought in as guest speakers. The seminar does not directly advance the work of the project, but it serves as a forum for reporting the work which has been accomplished and establishing a shared understanding of the project. It is also, during the school year, a weekly ritual which in itself provides continuity and cohesion.

Much of the actual work is conducted face-to-face and via electronic mail. There are several work groups – a visualization group, a network group, and text retrieval group, and others. Some of the individuals in the groups are at the same or neighboring locations, and each group has one or more electronic distribution lists.

Twice a year, all the active researchers, graduate students, corporate representatives, and other participants are brought together for retreats. During this time group meetings and individual discussions extend participants' familiarity with each other and each other's work. Time is planned for unstructured exchange in addition to formal meetings.

**Proprietorship and data integrity.** The sharing of data and analytic routines has not yet surfaced as a potential problem. Global change research is already highly collaborative in nature, and so far the attitude is one of delight that Sequoia will allow even more collaboration.

**Technology transfer within the project.** The global change researchers play a crucial role as testers and users of the technology. Introduction of a new technology has impacts on the receiver that are technical, institutional, economic, and personnel-related. Technical impacts can include the need for the receiver to gain expertise in use of the new technology, a need for some supporting technology which the receiver may not already have, and adjustments in the shift from one technology to another. Institutional impacts may include such things as the need to provide additional space and establishment of new policy structures. Impacts on personnel can involve the need to hire additional people for maintenance, possible shifts in responsibility as use patterns evolve. And all of these, in turn, can have economic consequences.

## 5 Conclusion

Cooperation between universities and industry have, according to Baba (1987), historically occurred primarily during periods of rapid technological change or intense international competition or crisis. Both of those forces characterize today's economy, and are driving the trend for increasing cooperation. Sequoia 2000 was designed to embody the critical issues in collaboration, technology transfer, and communication which must be solved if cooperative work is to succeed. Study of the problems and solutions created within Sequoia has the potential to assist future collaborative planners and to increase understanding of the relationships between information transfer and the growth of knowledge.

## References Cited


Auster, E.R. (1990). The interorganizational environment. In F. Williams and D.V. Gibson (Eds.), *Technology Transfer*. Newbury Park: Sage Publications, pp. 63-89.

Baba, M.L. (1987). University innovation to promote economic growth and university/industry relations. In P.A. Abetti, C.W. LeMaistre, R.W. Smilor, and W.A. Wallace (Eds.), *Technological Innovation and Economic Growth: The Role of Industry, Small Business Entrepreneurship, Venture Capital, and Universities.* Austin: $IC^2$ Institute, University of Texas at Austin.

Irwin, H. and More, E. (1991). Technology transfer and communication: lessons from Silicon Valley, Route 128, Carolina's Research Triangle and hi-tech Texas. *Journal of Information Science*, 17, 273-280.

Johnson, E.C. and Tornatzky, L.G. (1981). Academia and industrial innovation. In G. Gold (Ed.), *New Directions for Experiential Learning: Business and Higher Education – Towards New Alliances*, no. 13. San Francisco: Jossey-Bass, pp. 47-63.

Kraut, R.E., Egido, C., Galegher, J. (1990). Patterns of contact and communication in scientific research collaboration. In J. Galegher, R.E. Kraut, and C. Egido (Eds.), *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Larsen, J.K. and Wigand, R.T. (1987). Industry-university technology transfer in microelectronics. *Policy Studies Review*, 6, 584-595.

McHenry, K.W. (1990). Five myths of industry/university cooperative research – and the realities. *Research - Technology Management*, 33, 40-42.

Werner, J. and Bremer, J. (1991). Hard lessons in cooperative research. *Issues in Science and Technology*, 7, 44-49.

Wigand, R.T. (1990). University and microelectronics industry: The Phoenix, Arizona, Study. In F. Williams and D.V. Gibson (Eds.), *Technology Transfer*. Newbury Park: Sage Publications, pp. 132-152.