Chapter 6 --- Conclusions and Prospects for Research Using Research Webs

This dissertation has presented several major topics: a critique of the current research environment as it relates to the conduct of large-scale, long-term collaborative research; a design for a new approach to such research; a set of tools designed and implemented for this new research approach; and empirical studies of both the Research Web concept and on of its major tools, DocReview. These topics have built a progressive argument for their adoption by research teams that meet certain standards of team makeup and issue domain properties. This chapter makes the case for the adequacy of the argument, and continues on to demonstrate through rhetorical questions how successful social science research can be conducted using Research Webs. The Chapter concludes with a challenge to potential conveners of Research Webs.

6.1 Contribution to Knowledge

This research has described a new concept called the Research Web (RW) that is designed to improve the efficacy of long-term large-scale collaborative research in the social sciences. The RW builds on concepts and techniques familiar to all researchers, but augments those existing concepts and techniques by appropriating modern information technology, specifically the Internet and WWW. The RW assumes an environment that is collaborative and a team that is dispersed. The design basis of the concept is found in a very harmonious merging of modeling, as advocated by scientific realism[1], a research methodology -- VNS[2]; and a suite of information tools that support the functions of conventional scholarship. The RW is firmly grounded in arguments presented in Chapter 2. This grounding produces a synthetic creation that provides scholarly backing from numerous disciplines including, among others: philosophy, psychology, economics, sociology, geography, and information technology.

In order to fulfill the promise of the RW concept, information tools were invented or appropriated to serve dispersed collaborative teams. Existing Internet programs provide the means of communication, storage and distribution. Invented programs provide new channels for collaborative interaction, a scholarly apparatus, and a critical apparatus. Collaborative interaction in the development of a glossary for the issue domain is provided by Lexicon (see §4.8), which populates the Annotated HyperGlossary (see §4.5). Bibliographic information is assembled into an annotatable format in the Annotated HyperBibliography (see §4.4). The scholarly apparatus of the RW is a hyperdocument format called the Research Web Essay that integrates the textual content with an enhanced bibliographic service, language definitions, improved graphics, and hypertext links to supplemental information. The critical apparatus of the RW is a tool called DocReview (see §4.3) that allows any document to be presented in an annotatable format. Every Research Web Essay contains a link to a DocReview of its own content.

A Research Web will begin with a well thought out definition of the issue domain and a long list of interdependent research topics that will contribute to a comprehensive knowledge of the issue domain. In the preproposal phase, interested scholars will be recruited from the appropriate disciplines to collaborate in the examination of those research topics. The research topics will support a number of authoring teams all working on interdependent topics. After the proposal is approved, the work begins and the authoring teams begin to codify their knowledge into permanent models of parts of the issue domain. Since the research topics will be interdependent, researchers from every authoring team will have an interest in most other topics. With the understanding that collaboration is not merely lip service to some ideal set by funding agencies, all members should become active participants. Leadership and peer pressure, if not actual penalties, will ensure compliance with the injunction to collaborate.

Authoring teams will model the objects and processes observed in their research topic area. The models of the issue domain will accumulate knowledge from every authoring

team and all the reviewers of documents. Since all the research topics are interdependent, there will be a large number of common objects and processes. Each team will contribute nuances to each of the modeled objects. With the help of the facilitator, the models are linked to Research Web Essays written by the authoring team and to other documents. The Essays will form the basis for research reports, thus linking the RW and the authoring teams to the existing reward system of science.

Criticism will become the engine for successive refinement of the knowledgebase of the team, especially models and Essays. The knowledgebase will be an ever-expanding source of hypotheses suitable for providing the intellectual capital for further research proposals. The RW's leadership and committed scholars will likely issue a continuing stream of proposals in order to maintain a set of authoring teams that is large enough to maintain critical mass. As the body of literature produced by the team grows, the knowledgebase accumulates, and the reputation of the team becomes established, the proposals may become more attractive to the granting agencies.

Since the RW has a large number of participants, the loss of some by retirement, loss of interest or death, can be balanced by recruitment of new scholars. Should the issue domain become known territory, the issue domain may migrate to an issue adjacent to the original issue domain. Since the team will have been contributing to a context model that maps the boundaries of the issue domain and connections to adjacent issues, such a migration will be easy to make.

The Research Web and its associated tools were placed into actual service in numerous places by individuals and teams. Often the tools were used to augment conventional scholarship or to perform management functions. There were a few attempts to build functioning Research Webs. We discuss those attempts next.

6.2 Conclusions from Empirical Findings About Research Webs and Tools

Based on the literature and empirical results found by studying three attempts to establish Research Webs, there were three major problems: critical mass, funding, and selection of a suitable issue domain. Scholarly evidence was found to support the importance of critical mass[3,4]. Based on the relative success of the three case studies of RWs, critical mass was approached at 18 researchers in the Chromium VI RW. There was a decline in viability with declining team size in the three cases. The literature gives several examples of the success of large projects[5,6,7,8] that clearly reached critical mass. The size of those projects all exceeded the size of the most successful RW.

Funding was sufficient in two of the three cases, but one RW, the Soil Crust RW was damaged by lack of funding. After the single funded project was completed, that RW failed. Adequate funding is funding sufficient to support each of the authoring teams and their share of the overhead required to support the RW. This fact points out the need to include a share of the RW overhead in every research proposal. Justifying that share in the budget should not be difficult, and the requested overhead funding may in fact give the granting agency some confidence that the work will be managed as part of a larger research effort. The RW's founding proposal faces the special burden of having to ask for overhead funding that presumes that the overhead will in short order be shared by other research grants. This founding proposal is similar to a proposal to found an Institute, and such a proposal is in itself an expensive and long-term task[9].

Selection of a suitable issue domain is an intellectual problem that must be faced by the management of the RW. There are limiting factors in defining the issue domain: defining a scope that is both large enough to develop critical mass, and small enough to ensure that the topics of the authoring teams will be interdependent. In two cases the RW concept was applied to preexisting teams with either inappropriate or ill-defined issue domains. Attempting to establish a RW based on an existing research team assumes that the scope of the issue domain just happens to be appropriate. Establishing a RW on a

small grant with a limited scope assumes that grants can be obtained for closely related research problems -- growing by accretion or bootstrapping. Both of these strategies are very risky, betting on the hope that something will develop.

Case studies of DocReview (see §5.1) showed that the nature of dialog produced in the interaction of the team with documents was of a different character than the dialog seen in face-to-face meetings (see §5.1.7). This dialog was less inclined to be emotional, and was more inclined to be constructive; in other words it was a vehicle that fostered reflective knowledge building. DocReview was enthusiastically accepted and used by not only RW teams, but by many other users for widely varying documents. It was used to form the backbone of initial drafts of research papers, was widely used to review meeting minutes, was adapted to AutoCAD (a drawing and mapping program) to allow annotation of civil engineering plans, and provided a vehicle for evaluating user comments in a participatory design exercise. In short, DocReview was not only successful in its intended use, but was flexible enough to be adapted to tasks outside its intended purpose. Software products that followed its introduction were quite similar to DocReview, validating its design.

The Research Web Essay incorporates features that vastly improve the information delivery of scholarly documents. There is nothing that conventional research reports have that the Essays do not have, aside from being printed on paper. If an Essay is printed out then it has even more features than most printed reports (a glossary). The ability to call up bibliographic information, sidebars, and glossary definitions at a click represents a service that is just now becoming more widely seen with the rapid adoption of electronic journals. Research Web Essays are superior to every electronic journal I have seen, without exception. This superiority is due not only to the features mentioned, but to the integration of the critical apparatus of DocReview, thus allowing controlled peer annotation of the reports.

The Annotated HyperBibliography has drawn attention from the community of scholars assembled in discussion forums on the Internet. Though impressed and interested, none have found a place to apply it, or have been willing to invest the time to set it up.

Tools were developed for the facilitator to ease the creation of DocReview installations, the Annotated HyperBibliography, the Annotated HyperGlossary, and for the creation of Research Web Essays. Though the users of the Research Web never see these tools, they are critical to the enterprise. Without these tools the burden of creation of the web pages that embody these tools would be so great that their production would simply not be practical. The tools allow the facilitator to create the web pages with form-driven scripts that minimize the information that needs to be provided to the computer that does the actual page layout.

6.3 Prospects for Research Using the Research Web Contribution to Knowledge

Below I pose a series of questions that help the reader focus on the contribution to knowledge made by Research Webs and the tools that implement them. These questions frame arguments that confirm the utility of the dissertation's contribution to knowledge.

Can the Research Web effectively host large-scale social science research?

I described modern research problems in the social sciences as being more difficult than the problems of the past. These problems are of such complexity that large interdisciplinary teams are needed to attack them. Such a set of problems can be examined effectively only by a collaborative, long-term, large-scale team. If such a team were engaged in building the knowledge to illuminate the issue domain, what environment would serve the team best? Certainly not a decentralized environment that allowed research groups to develop hypotheses independently and then design experiments and write up reports, leaving the job of linking their work to the issue domain to senior scholars who would write summary review papers to publish in the

flagship journals of their discipline. Yet this is exactly how conventional scholarship works today.

Can the Research Web serve the social science research community?

Social science progresses through the processes of description and argumentation in the literature. At this point in history, the literature is a body of publications recorded in books and an immense number of journals, some highly respected, others highly suspect. There are several units of scientific knowledge: the book, the research paper, reviews, brevia, letters to the editor, reports and other member of a broad class of poorly indexed literature called "gray literature." To that literature, add the WWW, popular press and other unreviewed documents.

The Research Web's team has access to all of this material, just as do conventional teams. In the RW, however, the literature is cataloged in a corporate bibliography, the Annotated HyperBibliography (AHB), in a format that can be annotated by any member of the research team. This cataloging adds an additional layer of peer review. In addition to the core literature, the RW team, because of its interdisciplinary nature, captures literature from disciplines related to, but outside the "home" disciplines of the research topics.

Part of the knowledgebase of the RW is new synthetic knowledge produced by its members. This knowledge is indexed in several ways and can be searched by chaining down the hierarchical models of the RW -- or can be "full text" searched. The models of the RW are treated just like many of the other documents: they are hyperlinked to related documents, and are annotatable by the team members. In the social sciences, there are few models that extend beyond the topic of a single research paper. There are, within each discipline, broad general models, text based "laws", and codified behaviors; but these are seldom strongly linked to the work within a research paper. The RW is charged with the responsibility of modeling the entire issue domain.

The language of dialog takes on very nuanced meaning within each discipline. Those outside the discipline will not have good knowledge of all the terms used. In the RW, the language of the issue domain may span several disciplines. To bring order into the language of dialog, the RW employs the Annotated HyperGlossary (AHG). In the AHG, scholars within each discipline may contribute definitions of each term. If the difference between meanings is great, then multiple meanings can be listed. For minor differences, the terms may be annotated or glossed to explain any disciplinary nuance. In this way the RW attempts to solve the language problem; in a conventional work, the language is the language of the "home discipline."

Distribution of the knowledge products of the RW can follow normal academic channels in order to establish rewards for the contributors. In addition to publication in conventional journals, the RW can distribute the same report, without charge, in the Research Web Essay format. That format augments the paper with much ancillary information, such as bibliographic information, definitions, and sidebars. This information can be displayed at a click of the mouse. And above all, the Research Web Essay can be annotated online by anyone, thus expanding the reach of the knowledge-building team to the world.

It must be noted that the creative side of science is frequently cited as "the reason" why conferences and conversations are essential to doing science; and that creative scientific activity can only suffer from departures from face-to-face dialog. The environment of the RW does not proscribe synchronous dialog! It does, however, ask that the participants of such meetings be responsible to their colleagues by providing a record of such proceedings. Meeting minutes should be published online, and Memoranda of Communications summarizing important points be reported to the team by e-mail.

So, can social science be properly served by the Research Web? Most definitely, but primarily within well-defined issue domains. Social science, painted broadly, cannot be

served by the Research Web – that is still the province of the philosopher and senior scientist. The RW will, however, be a canonical reference for knowledge within its issue domain. If someone doesn't like what is there, they can annotate it!

Can the Research Web foster collaboration?

The RW is designed for collaboration. Indeed, people who are disinclined to collaborate are likely to be asked to leave. The objective of the RW is to develop comprehensive knowledge of the issue domain. That comprehensive knowledge is built by the team through collaborative participation in many activities, including model-building, criticism of documents, assembling knowledge from existing sources, and writing synthetic works. The number of channels or work objects available to join in the collaboration is much greater in the RW than in conventional teams.

Collaboration within the RW can actually be measured objectively. Annotation to documents and writing memoranda of communication may be measured by a simple word count. Under the assumption that all words written are actually contributions to the knowledgebase, those who write more annotations and report more ideas are better collaborators. The discovery of literature, model building, and other activities not directly associated with the development of a research report must be measured more subjectively. Clearly those who do not engage in such activities contribute nothing to them.

The recommended practice of acknowledgement of anyone who contributes to an Essay is an incentive to collaborate. Though acknowledgement is an undervalued practice in conventional research, it is today more widely practiced in some disciplines than in the past[10]. All annotations and e-mail contain the author's name, for attribution if not acknowledgement.

The emphasis that the RW places on collaboration will attract people who feel comfortable with cooperation; and conversely, may alienate people who are of a more

competitive nature. The conveners will have to make it very clear to scholars being recruited that collaboration is expected and that building a knowledgebase for the issue domain is the proximate goal, though the ultimate goal must remain publication of research reports. It may be a difficult adjustment for some people to understand that building a knowledgebase is an accumulation of small pieces rather than building grand creations. Evidence of a person's collaborative nature might be seen as a record of contributing reports with multi-author teams.

Can the Research Web accommodate large groups?

The limits to the size of the team are set by the number of active authoring teams working on the topics suggested by the team. Each of the topics must lie within the issue domain, so the limit is the number of identifiable topics in the issue domain. Practically, research reports on each topic must meet the test of minimum publishable unit; so the number of topics is not infinite. The lower limit is set by the need to maintain a critical mass of active researchers. If there is an upper limit, it is set by need to maintain interdependence of topics. In an extremely fecund issue domain, some planning may be necessary in order to develop knowledge in a manner that maintains interdependence through time. Maintenance of interdependence through coordination of authoring teams insures that the research team does not fragment.

Fragmentation into disciplinary subgroups is an ever-present danger. Left to their own devices, researchers will naturally be attracted to the rewards of working exclusively in their "home discipline." Just such a fragmentation was seen shortly after the establishment of the Consortium for Risk Evaluation with Stakeholder Participation (CRESP) project. While the proposal called for interdisciplinary study of the problem of the environmental cleanup of the America's WWII and Cold War nuclear production facilities, in fact interdisciplinary study was never entered successfully. Management divided the research team into task groups that were dominated by single disciplines, rather than dividing the team into task groups that investigated problems that could

benefit from research viewed from several disciplinary viewpoints. The Chromium VI Research Web was a happy, but short-lived, exception to that management choice. If each of these problems had been components of a process that contributed to an understanding of the overall issue of cleanup, then the initial objectives of the project might have been met.

Fragmentation by self-segregation of geographic clusters of researchers is also a threat to the unity of the research team. The pressures to revert to face-to-face communication (media competition) lead to a tendency to avoid documentation and thus knowledge sharing. When the MacArthur Foundation reorganized their Research Networks from nodes of collocated teams to an organization that selected members based solely on their ability to perform and to work in a collaborative environment, the Networks began to work much better[11].

Can the Research Web survive the passage of time?

There are two problems maintaining a long-term research project: maintaining interest and withstanding personnel turnover. If the Research Web produces new hypotheses at the rate expected, then there should be little difficulty in maintaining interest. The RW can be no less productive than conventional teams because all the mechanisms for hypothesis generation in conventional teams are present in the RW as well. In addition to conventional hypothesis generation, the RW has the added benefit of having more channels to carry follow-on activity. The models, for instance, will contain information on objects and processes that may not have been totally elaborated. Those partially defined elements may form the basis for additional investigations. Models will also have benefited from the attention of scholars from more than one discipline, and thus be richer than models developed conventionally.

The RW has within its knowledgebase not only models and essays, but also a body of e-mail and annotation, both of which may contain conjectures and paths for future research.

The third phase of research in the VNS is a building of robustness that will mine all sources for means of corroboration of results. Corroborating studies will enrich models further as well as providing opportunities to publish more research reports.

The problem of turnover may be reduced to some extent by the practice of providing graduate students and junior scholars with the opportunity to join more senior researchers on authoring teams. This "legitimate peripheral participation[12]" is an important practice and should be encouraged within the RW. Giving beginning scholars the opportunity to join established scholars in actual research will not only socialize them into the community, but should also develop them into skillful collaborators who are intimately familiar with the issue domain. One hidden benefit of the RW is its production of trained collaborators.

Recruitment of new members or of entire teams is an ongoing management task. The RW's public partition can be used to help in recruitment by introducing anyone who comes to the site to the environment that the research team works in. If a scholar is closely associated with one of the authoring team's research topic, that scholar might be invited to participate in review documents placed in the guest partition (see §3.5.1).

Long life can be assured by a steady stream of research grants. The RW may be large enough to enjoy economies of scale in proposal production. Since there will be several experienced proposal writers in the research team (all the conveners and most of the lead authors), a proposal authoring team might be constituted in order to produce that steady stream of proposals. With a dedicated facilitator and probably an administrative assistant, such a team could apply the tools and techniques of the RW to the proposal production task.

6.4 Prospects for Research About Research Webs

While we wait for Research Webs to become established we can prepare for the opportunity to study them by designing more refined protocols. Due to the *post hoc* nature of the case studies of the three Research Webs studied in this work, some important data escaped recording. For instance, efforts made to recruit participation in DocReviews needs to be recorded in order to manage DocReviews more effectively. Recruitment of team members was not studied; as the three cases studied all came with established teams.

A more thorough study of collaborations in web-based scholarly communities may lead to insights into survival characteristics and effectiveness in accumulation of knowledge. As noted in §1.3, there may be many intranets that demonstrate some of the features of the Research Web.

Research into the nature of issue domains suitable for Research Webs needs to go forward. The relationship between issue domain scope and Research Web viability studied in §5.2.5.1 must be examined in more detail. The entire issue of how to define, delineate and contextually model the issue domain must be clarified for the use of potential conveners. Establishment of a Research Web by accretion needs to be studied. There are many examples of informative web sites[13] that attempt to serve their communities by recruiting members to contribute scholarly content. Perhaps some of these sites might be converted to Research Webs, or could establish RWs to guide research in some constituent issue domains.

The tools developed for the RW have applications in other areas. DocReview has been especially flexible. Improvements can certainly be made in the tools.   It may be possible to capture the elapsed time of a DocReview session by recording the ending time (the start time is already recorded). PicReview (§4.9.1) is an essential tool. It must be implemented and examined in applications, wherever they may occur.

6.5 Concluding Remarks

This dissertation is an argument for the establishment of a Research Web. The concept of the RW has been tested as a prototype and appears to be sound. However, the risks are substantial. The forces of inertia are extremely powerful and will not be overcome by those who chose research problems that fit well into the current environment: problems geared to a short term research cycle, the disciplinary cores, and the accumulation of minor academic credits. We must wait for the proper conjunction of research problems of the proper scope and conveners willing to take a risk for large payoffs. The issue is likely to be one that has failed to yield to conventional approaches, one that is beyond small-scale research, but is important, difficult, interdisciplinary, and has been explored by a few intellectual pioneers and is now ready to be examined in depth. The risks are the potential waste of building a large knowledgebase that few use. The potential rewards are comprehensive knowledge of the issue domain with a body of new canonical literature that will establish the research team as the commanding authorities.

We infer from the literature, the attempts to establish large collaborations, and from the case studies herein, that the concept is sound and the supporting tools developed and methods suggested support the concept. Proof of concept will emerge only with the success of a well-staffed, well-funded, committed research team lead by strong management able to recruit and sustain a collaborative team for a long time. That effort will be a case study of a proper size, one that has the characteristics of size, scope, staff, and leadership that have been discussed above. An expensive experiment to be sure, and not for the faint-of-heart. Should a team of conveners rise to the challenge, I offer to advise, assist in proposal preparation, give technical assistance, and will donate the software that I've developed.

Notes to Chapter 6

[1] Aronson, Harré and Way 1995

[2] Brinberg and McGrath 1985

[3] Markus 1987

[4] Marwell and Oliver 1993

[5] Bero and Rennie 1995

[6] MacArthur Foundation 2000

[7] Schade and Putlitz 1996

[8] Ideker *et.al*. 2001

[9] Task Force on Enhancing the Research Environment 2001, 11

[10] Cronin 2001

[11] Kahn 1993, 11

[12] Lave and Wenger 1991

[13] See the NetWatch column by Mitch Leslie in Science archived at www.sciencemag.org/netwatch.