A STUDY OF THE RELATIONSHIP OF COMMUNICATION TECHNOLOGY CONFIGURATIONS IN VIRTUAL RESEARCH ENVIRONMENTS AND EFFECTIVENESS OF COLLABORATIVE RESEARCH

A Dissertation

by

IFTEKHAR AHMED

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2009

Major Subject: Communication

A STUDY OF THE RELATIONSHIP OF COMMUNICATION TECHNOLOGY CONFIGURATIONS IN VIRTUAL RESEARCH ENVIRONMENTS AND EFFECTIVENESS OF COLLABORATIVE RESEARCH

A Dissertation

by

IFTEKHAR AHMED

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Marshall Scott Poole Michael T. Stephenson
Committee Members,	Richard L. Street, Jr. Evan Anderson
Head of Department,	Richard L. Street, Jr.

August 2009

Major Subject: Communication

ABSTRACT

A Study of the Relationship of Communication Technology Configurations in Virtual
Research Environments and Effectiveness of Collaborative Research. (August 2009)
Iftekhar Ahmed, M.A. University of Dhaka; M.A., West Texas A&M University
Co-Chairs of Advisory Committee: Dr. Marshall Scott Poole
Dr. Michael T. Stephenson

Virtual Research Environments (VRE) are electronic meeting places for interaction among scientists created by combining software tools and computer networking. Virtual teams are enjoying increased importance in the conduct of scientific research because of the rising cost of traditional scientific scholarly communication, the growing importance of shared academic research by geographically dispersed scientific teams, and changes in the corporate research structures. New facilities provided by the Internet technology enhanced this situation. Currently, our knowledge about VRE-based scientific communication and what makes it effective is relatively immature in terms of understanding technology (interface, architecture, and software evaluation), system management (software systems, visualization, scalability), knowledge bases, expert systems, and coordination. Moreover, we do not have a comprehensive classification scheme for virtual research environments primarily from a technological viewpoint.

This study provided an analysis of VRE from a technological standpoint and developed a conceptual model that identified factors facilitating collaboration effectiveness with a primary focus on technology. VRE portals were at the core of the investigation as they are the entry points for VRE related information and resource access. First, the study developed a methodological framework for characterizing VREs, applied that framework to examine and classify existing VRE systems, and developed a new classification. Then, the study established a relationship between the technological profiles of various types of VREs and their productivity. Study results show that the technological arrangements of the VRE neither depend upon scientific discipline nor the existing functional typology. The study did not identify a significant presence of communication and collaboration technologies within the VRE systems. However, results indicated that there were a correlation between communication and collaboration technologies and VRE effectiveness.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION: THE IMPORTANCE OF RESEARCH

Virtual teams are enjoying increased importance in the conduct of scientific research. The rising cost of traditional scientific scholarly communication coupled with the facilities provided by the Internet (Esler & Nelson, 1998), the growing importance of shared research by geographically dispersed teams in different disciplines, and changes in organizational structure, especially in research and development wings of corporations, are transforming traditional scientific communication and research practices.

Virtual Research Environments (VRE) are electronic meeting places for interaction among scientists created by combining software tools and computer networking. In the scholarly literature the term collaboratory is often used as a synonym for VRE. A “space” in which scientists dispersed across different geographic locations work together has been termed a “collaboratory” (Kouzes, Myers, & Wulf, 1996). William Wolf of the University of Virginia coined the word collaboratory in 1989 (Kouzes, n.d.). In his definition, collaboratory is: a “‘center without walls’, in which the nation’s researchers can perform their research without regard to geographical location -

---

This dissertation follows the style of *Human Communication Research*.

interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries" (Kouzes et al., 1996, p. 40). Cyberenvironment and Virtual Laboratory are the other terms often used in the literature as a synonym to VRE or collaboratory. However, VRE is the most generic term used across the continents and represents the core idea of a virtual environment facilitating research activities, so that is the term we will employ in this study.

The VRE provides the possibility of improved coordination and collaboration among geographically dispersed scientists by enhancing communication and facilitating access to information. This possibility faces a limitation posed by the participating scientists' willingness to collaborate and coordinate their work. The group and organizational communication literatures discuss different aspects of collaboration and group effectiveness (Barge, 2002; DeSanctis, D'Onofrio, Sambamurthy, & Poole, 1989; DeSanctis & Poole, 1994; Moreland & Levine, 1988; Poole & Baldwin, 1996; Poole & Roth, 1989). Collaboration generally refers to the process of working together as a team or a group to achieve a common goal. Collaborative effectiveness, therefore, is the degree of effectiveness as a team in achieving that goal.

Because of the logical framework of the scientific problem solving process, characteristics of information and data, distinct types of knowledge and expertise of group members, and other communication, personal and organizational factors, teams that work with scientific discovery and/or research do not easily fit within the categories of traditional typologies of teams. Consequently, the collaboration process and collaboration effectiveness are unique in scientific groups compared to other types of groups.

A review of the literature suggests that knowledge about VRE-based scientific communication and what makes it effective is relatively sparse. In 1996, Rice and Boisvert reported that our knowledge of the virtual problem solving environments (PSE) was immature in terms of understanding technology (interface, architecture, software evaluation), in understanding system management (software systems, visualization, scalability), and in understanding knowledge bases, expert systems, and coordination. After ten years, scholars are still reporting similar limitations in our knowledge base and conducting studies to overcome these limitations (Benford, Greenhalgh, Rodden, & Pycock, 2001; Fraser, 2005; Hey & Trefethen, 2005; Polys, Bowman, & North, 2004; Zhao & Georganas, 2001; Zhuge, 2005).

Present research on VREs is largely based on case studies of one or a few VREs that provide snapshots of project work, particular descriptions of effective projects, descriptions of present and emerging technologies, and discipline-based recommendations regarding coordination and collaboration. There have been fewer studies that consider larger samples of VREs or engage in comparative analysis (Bos, Zimmerman, Olson, Yew, Yerkie, Dahl, & Olson, (2007) and Kouzes et al. (1996) are exceptions). As a result there are several gaps in existing knowledge, which this dissertation attempts to address.

First, there are numerous VREs currently under development. Those VREs are investigating appropriate cyberinfrastructure for research facilitation, effectiveness of specific tools present in the cyberinfrastructure, and scientific group effectiveness in a virtual setup. Sakai (http://sakaiproject.org/portal) and JISC (http://www.jisc.ac.uk/) projects are examples of those activities. Sakai is collaboration among various academic, governmental, and commercial entities to develop collaborative technologies for virtual learning environments. The Joint Information Systems Committee (JISC) is a UK funded project that is experimenting with information and communication technologies to support education and research. Often there is little guidance in building these VREs regarding what technologies should be included and what combinations are most effective. Experimental VRE developments for various disciplines are an integral part of the project. Publications related to such developmental activities often provide project oriented reports.

Second, the actual number of existing VRE is relatively small and information about those VREs are not readily available or organized. This dissertation will develop a typology of VREs based on empirical measures that will suggest dimensions along which VREs vary and help provide an organizational scheme for VREs. As will be shown existing classification schemes tend to discuss VREs from a discipline-based functional perspective. This approach tends to overlook technological aspects. This dissertation attempts to contribute to the literature by studying a relatively large sample of VREs in detail while maintaining a key focus on technology.

Third, there is an absence of proper methodology to study research related cyberenvironments. VRE portals are the gateway to research related cyberenvironments. They are different from general web pages. One way to study VRE is through the portals as they are the primary access to information and other scientific resources. However, there is a lack of proper methodology to evaluate these portals. This dissertation will develop a method for studying and evaluating VREs in terms of their technological affordances for researchers.

One of the most important aspects of VREs is the nature of technology-mediated communication. Human communication or collaboration in VREs includes all the complexities of traditional collaboration. Moreover, virtuality provides additional complexity. VREs incorporate many different technologies. Hence, collaboration patterns encouraged by the technologies might be expected to differ. There is no question that effective collaboration is an essential component of successful research. In regards to virtual scientific collaboration, cyber-infrastructure that incorporates appropriate tools and technologies to facilitate communication and collaboration are the core areas with less mature knowledge. Measuring the relationship between VRE effectiveness and technology, thus, becomes complex.

Currently we do not have a comprehensive classification scheme for virtual research environments primarily from a technological viewpoint. The present classifications are based on the function of VRE (e.g., Bos et al.'s (2007) categories of distributed research center, shared instrument, community data system, etc.) or field affiliation. Those classifications consider only the functional aspects of technology (whether the technology is facilitating data dissemination or facilitating experiments or observations). Technology is at the core of cyber-infrastructure not only for research facilitation from a functional perspective. As we are talking about a virtual environment, communication and collaboration among scientists also depends upon technology. A comprehensive look at VREs should consider all the present technology facilitating different aspects of the research process. Therefore, there is a need for a more empirically-grounded study to identify technological configurations of VREs and variations within them. Metrics for classification of VREs would facilitate exploration of the relationship between their structure and effectiveness in promoting collaborative research.

The purpose of this study is to identify factors that influence that effectiveness of the scientific communication and collaboration process in virtual research environments. The study will shed light on the current status of virtual environment technology including communication, collaboration, data-management, visualization, and coordination technologies. This study will provide an analysis from a technological standpoint and try to develop a conceptual model that identifies factors facilitating collaboration effectiveness with a primary focus on technology. It will do so in two steps.

First, this study will try to classify a set of existing VRE systems based on their technology. A scheme to classify VREs will provide us with a structured view to study, analyze, and compare issues related to collaboration patterns, technology use, and technology fit. An online examination of technology will be conducted. As there is no widely accepted method to examine VREs online, the study will develop a methodological framework for characterizing VREs. Then the framework will be applied to examine and classify existing VRE systems. The objectives of the online examination of VREs are to (i) identify the technologies in use within each VRE system under investigation and to (ii) classify VREs based on the configuration of technologies they incorporate. This will result in an empirically derived typology of VREs which can be compared to existing conceptual typologies such as those advanced by Bos et al. (in press).

Second, the study will try to assess the effectiveness of the VRE types in promoting collaborative research. In order to do this we must first develop a framework to measure collaborative effectiveness. Then an analysis using the VRE typology and collaborative effectiveness framework will be conducted to investigate the relationship between (i) the technological profiles of various types of VREs and productivity and (ii) particular technology setups within each type and VRE productivity.

Chapter II of the study will review the existing literature related to groups and teams including virtual teams, the nature of scientific inquiry, available technologies for virtual collaboration, and existing VRE classification schemes. This review will help us understand the present situation by providing a summary of our present knowledge, connect different aspects of communication and information technology literature, and identify the lacks in knowledge and understanding. A rationale for research and specific research questions will be developed based on the discussion. Chapter III will introduce research methodology. The chapter will provide a detail description and arguments behind the proposed sample selection, framework for data collection, and analysis. Chapter IV will provide research findings. Chapter V will provide a comprehensive discussion of findings, and also implications and limitations of this research.

CHAPTER II

LITERATURE REVIEW

Concerns over the effectiveness of groups and teams have produced a vast amount of literature. We have experienced the development of a new line of discussion with the introduction of the Internet in the early 1990s. Since then, scholarly interest related to virtual teams has provided significant insights into the functional, social, and psychological processes of virtual groups and teams, group effectiveness, and the importance of technology within those environments.

Scientific knowledge discovery is a very distinctive process due to the nature of scientific information and data, characteristics of individuals related to the process, and the stages involved in the discovery process. We can identify different models describing the stages of scientific discovery. The traditional literature of the field is based on co-located scientific team processes. Formation of scientific activity related virtual teams is a comparatively new phenomenon. Consequently, the related knowledgebase is much weaker. However, we can identify scholarly attempts to explain this new phenomenon. Moreover, to face the challenges of the new century and also to work efficiently and innovatively, organizations are implementing virtual arrangements (Paré & Dubé, 1999). Computer mediated communication is changing the nature of work teams, reducing the need for co-workers to be co-located, and becoming an integral part of scientific work (Walsh & Bayma, 1996).

This chapter will, first, discuss the literature related to groups and teams, virtual teams, and the scientific discovery process. Then it will try to explain virtual scientific teams by combining our knowledge of the traditional scientific discovery with virtual teams. The discussion will establish the importance of technology in virtual scientific process by looking at the enabling and constraining factors of technology. Finally, a rationale to study VREs will be developed based on the discussion of present literatures.

Groups, Teams, and Virtual Teams

Before turning to scientific inquiry, we will first introduce some background on groups and teams. VREs are host to teams of scientists and hence some background on groups and what makes them effective will be useful in understanding VREs and what contributes to their effectiveness.

Groups and Teams

The similarity among the definitions of group and team suggests that the meaning of groups and teams often overlap. Some scholars believe that teams are more synergistic than groups. Katzenbach and Smith (1993), for example, assert that groups become teams when they develop a sense of shared commitment and strive for synergy among members. However, following Guzzo and Dickson (1996), this paper, while recognizing that there may be degrees of difference, will employ the labels “team” and “group” interchangeably.

Kozlowski and Ilgen (2006) define teams as

> (a) two or more individuals who (b) socially interact (face-to-face or, increasingly, virtually); (c) possess one or more common goals; (d) are

> brought together to perform organizationally relevant tasks; (e) exhibit interdependencies with respect to workflow, goals, and outcomes; (f) have different roles and responsibilities; and (g) are together embedded in an encompassing organizational system, with boundaries and linkages to the broader system context and task environment. (p. 79)

Kozlowski and Ilgen’s (2006) definition accepts the dyad as a team, as well as groups with larger sizes. This definition fits the nature of scientific collaboration of this study.

Recardo and Jolly (1997) distinguish four types of teams: simple problem-solving, task force, cross-functional teams, and work teams (see Figure 1). Another classification of teams by Thylefors, Persson, and Hellström (2005) based on a continuum of collaborative intensity (discriminated by six themes: role specialization, task interdependence, co-ordination, task specialization, leadership and role interdependence) among team members, places cross-professional teams in three major categories. From low to high on the collaborative continuum, these are multiprofessional, interprofessional, and transprofessional teams.

Scientific collaboration teams range across these types and continua. In different phases, scientific work might start with simple brainstorming and later move into complex data analysis. However, based on the nature of scientific discovery it would be troublesome to argue that a scientist is not deeply involved during the brainstorming phase. Based on the nature of research, sometimes initial idea development phase may require more involvement than data gathering phases. Again, the level of involvement may fluctuate during data analysis and report preparation phases. This scenario tells us that the level of involvement in scientific discovery process is nonlinear. Consequently, Recardo and Jolly's (1997) team definitions based on increasing involvement and impact fall short in classifying scientific teams. Similarly, in different parts of its work, a scientific team's members might work together or independently, in loosely coordinated and in highly structured environments. It is possible to discriminate different phases of a team's "life" using the collaborative continuum. However, it would be problematic to describe a scientific team using the traditional labels.

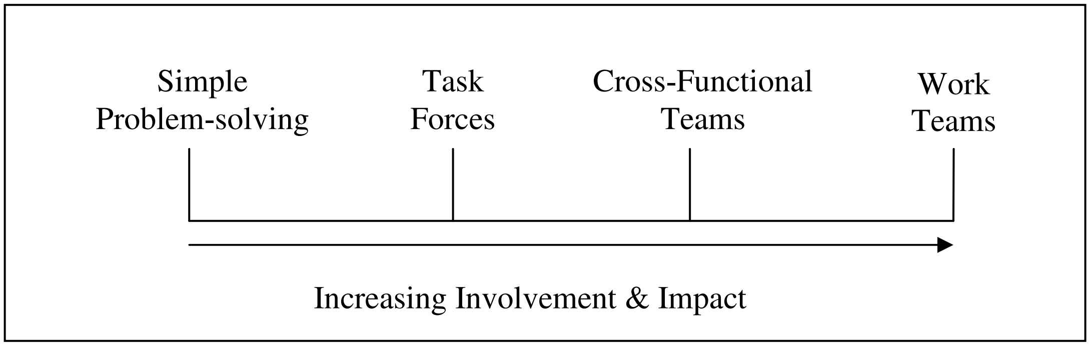


Figure 1 Recardo and Jolly's (1997) classification.

Due to this, traditional group effectiveness arguments are not wholly applicable to scientific productivity. The majority of early team effectiveness literature uses McGrath's input-process-output (IPO) model (Ilgen, Hollenbeck, Johnson, & Jundt, 2005). Translating scientific teams into McGrath's IPO framework, inputs refer to the composition of the team in terms of the constellation of individual characteristics and resources at multiple levels (individual, team, organization, etc.). Processes refer to activities that team members engage in, combining their resources to resolve (or fail to resolve) task demands. Output has three facets: (a) performance judged by relevant others external to the team; (b) satisfaction of team-member needs; and (c) group wellbeing (building the group as a team and increasing its functional abilities for the future) (Hackman, 1990). Processes mediate the translation of inputs to outcomes (Kozlowski & Ilgen, 2006). McGrath's model predicts that the group process - inter-group and intra-group actions that transform group resources into products - leads to group effectiveness. Hence, group effectiveness can be viewed as the major output of small group behavior (process or task). Group performance, satisfaction of members' needs, and the ability of the group to exist over time are the three components of group effectiveness (Gladstein, 1984).

Virtual Teams

The virtual team is an emerging form of group in the Information Society. The definition of virtual team depends upon member location and mode of communication. "Key defining features of virtual teams (VTs) are these: (1) Their members are dispersed and do not conduct much work face-to-face, and (2) most interaction between members is mediated by information and communication technologies (ICTs)" (Poole & Zhang, 2005, p.364). The complexity of virtual teams can be understood in terms of the nature of their task, geographical dispersion, team composition, diversity, nature of communication, and technology in use.

There are a variety of VTs ranging from simple social support groups on the net to multidisciplinary scientific collaborations. Task plays a critical role in virtual teams as the nature of task adds complexity because of its direct relationship to communication technology. Group activities that require more socio-emotional cues for group effectiveness tend to be more complex in the virtual environment than in face-to-face (FTF) settings. The nature of communication in VTs, depending upon task requirements and diversity factors, may require increased socioemotional communication (e.g.,, small talk in social support groups), structured communication (e.g., process flowcharts in software development groups), collaborative communication (free flowing with some restriction in decision-making groups), or specialized collaborative communication (as in scientific research groups that need a specific vocabulary).The location of the members of the VTs may range from different continents to different parts of the same building. Moreover, in many cases, several members of a group can be located in the same place and others in different geographical locations. When this occurs it divides groups into subgroups where each subgroup has a specific geographical location.

Virtual teams vary from zero-history groups to groups with complex composition where VTs may have several members new to the group and others with an established FTF social relationship. There are virtual teams organized around specific task requirements. On the other hand we experience naturally-emerging Internet groups. It is also possible to find naturally emerging task-based virtual teams. Different geographical locations and variation in expertise add a natural diversity to VTs. This diversity includes demographic factors, disciplinary factors, cultural and linguistic diversity, and temporal differences. ICTs enable and constrain communication in the virtual environment (Ellis, Gibbs, & Rein, 1991; MacEachren, 2001; Wilbur, 1997). Availability of required technology and adequate skills to use that technology can enhance group communication. The technology should match the nature of the task. Research findings show us that e-mail is very suitable in a hierarchical organizational communication where visual cues enhance socio-emotional group communication.

Virtual team technology ranges from simple technologies like e-mail, chat, BBS, and Weblogs, to complex group decision support systems (GDSS)/computer supported collaborative work (CSCW) technology. Poole and Zhang (2005) state that the available technology falls under one of four categories: (i) synchronous communication (telephone, teleconferencing, videoconferencing), (ii) electronic information sharing (e-mail, file sharing), (iii) metainformation on entries into the repository, and (iv) decision and process support.

For general purposes, Internet-based tools provide the primary technology for virtual team interaction. However, specific group activities related to decision-making, problem-solving, collaboration, and coordination require specific tools developed for the purpose. Most of these technologies are classified under the generic category of "groupware." Ellis, Gibbs, and Rein (1991, p.39) defined groupware as "computer-based systems that support groups of people engaged in a common task (or goal) and that provide an interface to a shared environment". Technologies similar to Lotus notes and GDSS are widely used among CSCW technology. Present collaborative technology often includes whiteboards, chat rooms, desktop videoconferencing, voting tools, and file sharing systems (Schur, Keating, Payne, Valdez, Yates, & Myers, n. d.). Based on physical proximity, GDSS sessions that facilitate problem-solving or decision-making can be classified under three categories: (i) FTF sessions where participants use a common place to use technology, (ii) distributed sessions where participants are in different locations, and (iii) mixed-mode sessions where some of the participants are in the same place and others are in different locations (Ellis et al., 1991).

Scientific group effectiveness can be viewed through the lens of these models of groups and virtual teams. However, as scientific discovery is the prime focus of scientific work, these models require modification. A discussion on scientific collaboration will clarify some required modifications.

The Nature of Scientific Inquiry

This section will consider two models of scientific inquiry that portray it as a series of phases of activity. Following this, we will suggest an integrated model of scientific work that can serve as a foundation for inquiry into VREs.

Lievrouw and Carley's Model of Scientific Communication

Lievrouw and Carley (1990) argue that "scientific activity can be viewed as a communication cycle having three progressive stages: conceptualization, documentation, and popularization" (p. 459). One major limitation of Lievrouw and Carley's (1990) scientific activity model is the absence of consideration of the actual research process. However, their discussion focuses on some important aspects of non-research processes.

In the conceptualization stage scientists explore a problem area, usually by informal interpersonal interaction. In this stage, participants share both scientific and social information. Lievrouw and Carley (1990) argue that communication in the conceptualization stage is primarily face-to-face (FTF) interaction among people who know and trust each other. The amount of social interaction decreases and scientific information increases in the next stage, the documentation stage. In this stage, discussions about concepts and methodologies help scientists negotiate their standpoints in the scientific paradigm, develop interest groups around various research topics, establish liaisons and gatekeepers for such groups, and produce a body of documented information. Creation and diffusion of new information occurs in the documentation stage. The activities of this stage are mostly facilitated by formal communication. The dissemination of the documented body of new information is the primary communicative activity of the final stage, popularization. Scientific ideas, through this stage, become a part of public discourse. The biggest challenge in this stage is to communicate the scientific data to a general public audience. Lievrouw and Carley (1990) define the general public as “individuals within the same society who share relatively little specific (scientific or social) information, but who do share a common culture” (p. 462).

Communicating scientific data to ordinary people helps a scientist expand his/her “sphere of influence,” which in turn enhances credibility. Successful communication is likely to expand the influence of scientists, includes more people, and covers a greater geographic area. However, the picture for scientific teams embedded in organizations is somewhat different. They need to communicate effectively not primarily for public recognition, but in order to satisfy non-scientist policy makers and administrators of these organizations. These managers, sometimes from external organizations related to financial and other resource allocation, influence scientific productivity through control of the human and material resources that the scientists require.

The nature of scientific communication, therefore, moves from informal to formal. The initial stages of communication are characterized by collaboration within a homogenous group setting that in later stages expands to heterogeneous groups and moves from small group communication to mass communication. Within initial stages, the communication process depends upon the exchange of a significant amount of scientific information. Hence, the effectiveness of group activities or discovery process largely depends upon effective scientific data and information management.

Between the conceptualization and documentation phases, scientists conduct research and go through rest of the documentation and popularization phases. These aspects are not considered in Lievrouw and Carley’s (1990) model. Considering the pattern of scientific discovery should suggest ways in which to remedy this gap.

A Model of Scientific Discovery as Problem-solving

The model of scientific discovery as problem solving fits well with existing work on group problem-solving and decision-making and thus promises to facilitate development of a model of scientific teams working in VREs. Scientific discoveries involve the process of hypothesis formation and hypothesis testing (Okada & Simon, 1997), which can be modeled as problem-solving.

Simon and Lea’s (1974) dual-space model asserts that a discovery process involves search in two problem spaces: a hypothesis space and an experiment space. “Hypothesis space search builds the structure of a hypothesis and uses prior knowledge or experimental outcomes to assign specific values to its features. Experiment space search tests hypotheses experimentally” (Okada & Simon, 1997, p. 110). Hypothesis space search uses strategies to search memory for possible hypothesis or experiments until new hypothesis can be generated from data (Dunbar, 1993).

Schunn and Klahr (1995) extended this two-space model into a four space model introducing two new spaces: the data representation space and the experimental paradigm space. Schunn and Klahr (1995) assert:

> In addition to search in an experiment space and a hypothesis space, scientific discovery involves search in two additional spaces: the space of data representations and the space of experimental paradigms. That is, discoveries often involve developing new terms and adding new features to descriptions of the data, and the also often involve developing new kinds of experimental procedures. (p. 1)

In their new framework, the hypothesis space of the dual processing model has been divided into a data representation space and a hypothesis space. Similarly, the experiment space has been divided into an experimental paradigm space and an experiment space. The data representation space deals with the representations or abstractions of the data chosen from the set of possible features. The hypothesis space deals with causal relations in the data drawn using the set of features in the current representation. The experimental paradigm space deals with a class of experiments which identifies the factors to vary and the components which are held constant. Finally, the experiment space deals with the parameter settings within the selected paradigm (Schunn & Klahr, 1995). Both the dual and the four-space model focus on the mechanics of scientific discovery rather than the cognitive or socio-emotional processes of the scientists.

The dual- and four-space models view scientific discovery as a problem solving process. Here, the problem consists of an initial state, a goal state, and a set of operators transforming the initial states to goal states. The set of states, operators, and the constraints on the operators construct a problem space. Scientific discovery comes through coordinated search in the problem spaces. The search for the path that links the initial state to goal state is the problem solving process. Effective problem solving consists of engaging in as few search paths as possible because problem spaces grow with every available alternative path (Klahr & Simon, 2001).

The discovery process may contain well-defined or weakly-defined states or operators. The definitions of states or operators are influenced by the particular problem and the available knowledge. Knowledge is of extreme importance in the discovery process as it engages cognitive processes and their role in the discovery (Klahr & Simon, 2001).

Knowledge and expertise in scientific discovery can be procedural or declarative. Declarative knowledge is reportable. In general, domain experts (experts in a particular area or discipline) have greater declarative knowledge than domain novices. However, a new scientific problem minimizes this distinction. For new problems experts and novices are about equal in terms of the problem specific knowledge they hold. However, the distinction stays firm in the background knowledge of the field or discipline (Schunn & Anderson, 1999). Procedural knowledge can be domain general (i.e., statistical analysis) or domain specific (i.e., molecular analysis). Those scientists who possess domain-specific procedural and declarative knowledge are called domain experts, while those who possess domain-specific procedural knowledge are referred to as task experts (Schunn & Anderson, 1999). The scientific collaboration process, beside scientists (domain experts) contains non-scientist domain specific task-experts and domain general task experts (i.e., engineers). Hagstrom (1964) states, “modern scientific techniques and instruments require skills not possessed by a single individual, and scientists often require the technical assistance of professionally trained persons” (p. 251).

However, one of the problems in the scientific discovery process also stems from this task division. It is often argued that the professional technicians, in many cases, are not motivated like scientists to engage in an innovative process. They work for money and other extrinsic rewards, not for recognition from the scientific community and are not expected to make research decisions or to show commitment to solve scientific problems (Hagstom, 1964).

Though the human factor influences the outcome of scientific discovery, the largest part of the effectiveness of the scientific process depends upon the logical framework of the problem solving process. Understanding the nature of scientific work, characteristics of information and data, and communication patterns thus become important.

During the process of scientific discovery, data analysis also follows a very specific and logical path. Springmeyer, Blattner, and Max (1992) describe the process of scientific data analysis based on two main activities: (i) investigation (exploring the data to extract information or to confirm results) and (ii) integration of insight (assimilation of the knowledge). Investigation activities involve interacting with representations, maneuvering, and applying mathematics. Maneuvering refers to tasks involved in organizing the data, and choosing and setting up representations. Integration activities involve maneuvering and expressing ideas. Applying math involves the derivation of mathematical quantities or the generation of new data using rough estimates to complex calculation and statistics. Their analysis describes the process of mathematics based sciences.

An Integrative Model of the Scientific Process

We can get a comprehensive picture of the scientific process by combining Lievrouw and Carley's (1990) scientific activity model with the four-space model. The stages of this model are defined in Table 3 below, along with criteria for judging how effectively the stage was carried out by scientific collaborators. We can argue that scientific discovery starts with conceptualization. The conceptualization phase incorporates the data representation and hypothesis formation stages. However, both stages will later initiate the documentation phase. Representation and hypothesis formation together provides a structured view of the initial state of the problem solving process. The experimental paradigm and experiment spaces provide the set of operators for the problem solving process. Investigation of the data analysis process links the representation state to the experimental paradigm. The documentation phase also involves documenting data. Later integration of insight stage of data analysis leads to popularization. There is also a process of task division and expert interaction. In all the phases the scientists and technicians must negotiate task division and how they will interact to maintain socio-emotional climate. (See Table 1 for an explanation of stages and processes).

Based on our analysis, we can argue that scientific group effectiveness depends upon the success of transforming an initial state into a goal state. In science, the problem at hand is not always well-defined. However, there is an interaction between the problem and the problem solver based on the process of problem recognition. Knowledge plays a critical role in this recognition. What kind of problem solving or search operators would be used to move from the initial to the goal state depends upon the definition of states and knowledge. Declarative knowledge (factual information about an event or a phenomenon) is often critical in problem recognition and to identify operators. Procedural knowledge (how to carry out a specific operation) becomes important in operating tools to move to later stages. Therefore, interaction between domain and task experts becomes important. During research, the data analysis process needs to support the move from one defined state to the other. This data analysis process often requires mathematical and statistical procedures to generate explainable and reportable results. Therefore, the process of transforming an initial state into a goal state depends upon the

Table 1 An Explanation of Stages and Processes

Terminology	What	How
Conceptualization	Scientists explore problem area, communicate ideas and interests with others, and share scientific and social information to formulate initial research and to develop mutual trust.	Discussions about concepts, methodologies, and style of communication. Negotiation of standpoints in the scientific paradigm. Development of research interest groups, liaisons and gatekeepers of such groups.
Problem solving process	A search for the path that links the initial state (problem to address) to goal state (discovery, solution, or greater understanding of the problem).	Managing paths. Depending upon the research problem, constraining paths to as few as possible because problem spaces grow with every available alternative path.

Table 1 (*Continued)*

Terminology	What	How
Hypothesis Space search	Using prior knowledge or experimental outcomes to build the structure of a hypothesis and to assign specific values to its features.	Develop strategies to search memory for possible hypothesis and to apply knowledge to develop hypotheses; Develop strategies to explore old data or to generate new data through experiments until a new hypothesis could be generated.
Investigative space search	Set up and conduct experiments, collect data through observation, or set up simulations.	Carrying experiments and other studies out successfully. Effectiveness depends upon how successfully the research moves from the initial state to the goal state.

Table 1 (*Continued)*

Terminology	What	How
Data analysis process	Exploring the data to extract information and assimilating that to the existing knowledge.	Successfully organizing the data, choosing and setting up representations, applying math, and generation of new data using rough estimates to complex calculation and statistics.
Documentation	Documenting initial idea discussions, hypothesis, experimental procedures and results, and data analysis and outcomes to produce a body of documented information that will direct the diffusion of new information.	Producing understandable (by the intended audience) and searchable documents.

Table 1 (*Continued)*

Terminology	What	How
Popularization	Dissemination of the documented body of new information to scientists and general public audience.	Disseminating various documented information to the intended audience. Effectiveness depends upon the understanding, recognition, and discussion of the reported information by the audience.

following factors: (i) defining states and operators, (ii), managing the nature of knowledge flow (iii) successful interaction between domain and task experts, (iv) following the logical path through the data analysis process, and (v) maneuvering and applying mathematics and other analytical procedures.

While the nature and sequencing of activities is one important aspect of science that VREs must support, also important is the content of scientific communication, the types of information they must handle. We turn to this in the following section.

The "Content" of Science: The Nature of Information and Communication during Scientific Inquiry

Scientific information is unique several respects. "Scientific information is logical information received in the process of cognition; it adequately reflects the phenomena and laws of nature, society, and thought and is used in social-historical practice" (Mikhailov, Chernyia, & Giliarevskii, 1984, p. 65). Mikhailov et al. (1984) present a typology of scientific information (with four basic divisions and each division having two sub-divisions) that clarifies some of its major characteristics (Table 2).

Because it was published before widespread use of the internet, this typology does not address web-based documents. Web documents are widely disseminated but do not fall under 'document' in classic sense. Web documents are different in that they can be expanded by linking them to explanations or additional information, in most of the cases are readily archivable, can be viewed simultaneously by many users, and can incorporate a search method that enable users to identify similar documents.

Scientific collaboration is often multidisciplinary in that it includes scientists and technologists of different disciplines. As a result, scientific collaboration requires both scientific and technological information. Scientific collaboration is different from much other organizational collaboration as there is a definite need for two distinct types of knowledge and expertise. Since technology involves applied science, technological information has both similarities to and differences from scientific information (Table 3).

A final piece of the puzzle of scientific inquiry is the nature of communication and collaboration among scientists. Several studies have focused on factors that promote effective scientific collaboration. They are discussed in the next section.

Table 2 Types of Scientific Information

Basis of Division	Named Sign	Application
Audience for information	Mass	Intended for everyone
	Special	Intended only for specialists
Type of transmitted information	Documental	In scientific documents
	Factographic	Transmitted ideas and facts, extracted from scientific documents
Medium of information	Published	Widely disseminated by means of documents
	Not published	Not considered for wide dissemination
Degree of analytical/synthetic information processing	Primary	Direct result of scientific research and experiment
	Secondary	Synthetic processing of primary information

Note: Adopted from Mikhailov et al. (1984, p. 70).

Table 3 Differences between Science and Technology as Types of Human Activity

Characteristics of information requirements	Science	Technology
Final Goal	Knowledge of laws of nature, society, and thought	Preparation of useful things
Motives	Human desire for knowledge	Satisfaction of societal needs
Societal Control	Weak	Strong
Timetable for Solution of Problems	Not established	Established
Nature of information Used	Noncomplex, Nonconcrete	Complex (multidisciplinary), concrete
Urgency involved in Answering Requests for Information	Small	Large
Preferred Method of Satisfying Information Requirements	Without interference from information workers	With help of information workers

Note: Adapted from Mikhailov et al. (1984, p. 119).

Scientific Team Communication and Collaboration Patterns

According to Lievrouw and Carley's model (1990), scientific communication starts during the conceptualization phase. The traditional framework for scientific collaboration is based on proximity. Proximity can be classified into two categories: organizational proximity (affiliated to same organization) and physical proximity among collaborators (Kraut, Egido, & Galegher, 1990) Physical proximity enables collaboration via meetings, problem discussion, planning, supervising, training coworkers, and other means and helps to avoid or to quickly address many problems related to research projects. Physical proximity helps scientists to choose collaborators based on an understanding of their intellectual capability and to identify potential partners even before starting a project (Kraut, Egido, & Galegher, 1990).Within the work environment, personal communication often affects task-related activities. Personal communication flows different ways in different disciplines. Research shows that "zoologists valued personal communication more highly than chemists and chemists more highly than biochemists" (Berelson & Sills, 1960, p. 51).

The collaborative research environment influenced by communication, personal, and organizational factors has significant impacts on scientific group performance (Pelz & Andrews, 1976). The importance of organizational and group influences is emphasized by Carley and Wendt's (1991) argument that the majority of the scientific research today is done by groups rather than by individual scientists.

Pelz and Andrews (1976) investigated factors that contribute to the development of a collaborative scientific research environment. They studied 1300 scientists and engineers from five industrial laboratories, five government laboratories, and seven departments of a university. They gathered data about the performance of the scientists and characteristics of climate where they work. A scientist's performance was measured by gathering data about

> his scientific or technical contribution to his field of knowledge in the past 5 years, as judged by panels of his colleagues; his overall usefulness to the organization, through either research or administration, also as judged by his colleagues; the number of professional papers he had published in the past 5 years (or, in case of an engineer, the number of his patents or patent applications); and the number of his unpublished reports in the same period. (p. xvii)

A questionnaire-based survey was employed to gather the climate data. According to Pelz and Andrews (1976) both freedom and coordination influenced the level of effectiveness in collaborative scientific work. They found that a loosely coordinated setting demands higher motivation from internal and external sources. Coordination settings refer to the level of autonomy present in the organization. Their 5-point scale ranged from very tight (highly coordinated with the presence of a structured organizational pyramid) to very loose (where individual enjoy greater autonomy and less supervision). A moderately loose situation works well with individual's autonomy. Pelz and Andrews (1976) reported that scientists with high level of interaction with colleagues were more successful in their work. In this interaction process, especially with decision makers, the persuasive ability of a scientist was reported to be associated with performance. Generally, scientists tend to be highly individualistic as higher education places more emphasis on individual accomplishments (Mohrman, Cohen, & Mohrman, 1995; Shannon, 1980). However, their self-directive work ethic, desire for challenging work, peer approval seeking, and interest in knowledge sharing make these individuals collaborative (Hackman, 1990).

Pelz and Andrews (1976) reported that higher performance was also positively correlated with task involvement. Though work satisfaction had a relationship with perceived contextual outcomes, it was found that performance had no significant relationship with disagreement among scientists on technical strategy or problem approaches. Moreover, a combination of technical disagreement and similarity in motivational sources (types of problems, career interest, and social relations) often resulted in high performance. Their research primarily focused on face-to-face work. However, with advances in information and communication technology, the nature of scientific collaboration has evolved to include more intense collaborative work among geographically dispersed colleagues.

During the initial stages of scientific communication when collaboration among scientists take place, geographical distance among members changes the nature of communication and technology-mediated communication becomes inevitable. As scientific progress depends on communication, whatever mode a scientific group chooses will influence the research process and its progress (Carley & Wendt, 1991).

Communicating at a distance differentiates scientific workgroups from conventional face-to-face (FTF) collaboration not only in terms of communicative medium but also in terms of the scope of extending research groups (Carley & Wendt, 1991). A new phenomena in scientific collaboration is the re-emergence of extended research groups defined as “a very large unified, cohesive, and highly cooperative research groups that are geographically dispersed yet coordinated as though they were at one location and under the direction of a single director” (p. 407). This re-emergence is based upon virtual team technology.

This opens the question of how VREs might support scientific inquiry by virtual scientific teams. In the next section we consider some of the technologies that might be used in VREs.

VREs and Technologies to Support Scientific Inquiry

This section will discuss the technologies that are available to support the scientific process in VREs and previous typologies of VREs. Ideally a VRE should support the phases of scientific activity, the exchange and processing of appropriate scientific information, and scientific collaboration, as portrayed in the previous three sections.

Numerous virtual scientific collaboration sites based on a variety of virtual team technologies can be found on the Internet. These sites mostly attempt to duplicate scientists’ traditional information sharing and co-authoring activities. However, a growing number of other scientific collaboration sites or collaboratories are appearing on the net. These sites go beyond the attempt to duplicate scientists’ traditional information sharing and co-authoring activities. Moreover, these provide a space for scientists to conduct collaborative research.

We know that VRE systems incorporate a number of different technologies. The missions of VREs differ and collaboration patterns encouraged by the technologies available also differ. Because of this complex variation, we currently do not have a proper classification for virtual research environments. In order to classify VREs, we need to consider some specific questions. These include: (i) the goals of a particular research environment, (ii) the nature of research that is taking place within the VRE, (iii) the technologies required for the intended research, and (iv) the technologies available within the environment.

Schur et al. (n.d.) identifies four types of collaborations that occur during mediated communication: (i) peer-to-peer, where scientists collaborate with colleagues through CMC devices that enable them to share instrument control, sketches, and raw data files without site visits, (ii) mentor-student, where mentors provide highly interactive lectures and training using previously prepared materials, (iii) interdisciplinary collaboration, where researchers with different background communicate summaries of experiments and results using commonly understandable terms, and (iv) producer-consumer collaboration, where researchers provide data as input to people with different background and goals.

Chin, Myers, and Hoyt (2002) identified five phases of scientific research group formation (associative, formative, explorative, active, and dormant) when they were exploring the transitions of role, duties, and expectations of the Virtual Nuclear Magnetic Resonance Facility (VNMRF) collaborators. In order to identify others' research interests, knowledge, expertise, and also to locate resources, scientists related to the VNMRF utilized both FTF and virtual communication (meetings, conferences, virtual gathering, newsgroups, mailing list, etc.) to develop wide-ranging networks of contacts. There were also formal relationships primarily to “facilitate access to instruments, sample distribution, and understand guidelines for instrument operations” (p. 89) which later transformed into an explorative relationship for sharing theoretical ideas, knowledge, and skills. This positive working environment gradually helped scientists to develop trust and comfort. Later it became an active collaboratory and fostered participatory relationships. However, “the loss of mutual trust, commitment, and sense of ownership sometimes halted active collaboration where the research project became indefinitely suspended” (p. 89). VNMRF scientific groups, therefore, showed both the possibilities and constraints of scientific group development.

Different collaboratory examples provide us with different pictures of collaboration. Most of these are based on specific projects. We still do not have a holistic picture of virtual scientific collaboration. We have ideas of what is happening in VREs, but we still do not know how that collaboration is taking place. However, we know that technology plays a significant role in the VRE effectiveness. Understanding the role of technology is not only important in understanding the ongoing collaboration process, it is critical to understand the limits of the possibility of virtual collaboration.

Though futuristic technologies show us some positive scenarios, numerous technological complexities and limitations arise when we analyze current VRE systems. Recent research illustrates ongoing arguments on technological issues related to the requirements to support real-time interactions, graphical and behavioral complexity, system bottlenecks created by network traffic, local computers' capability to process and render information while maintaining standards, and audience-centered operations including flexible and dynamic interest management schemes, distributed architecture, principles of sharing and visualization, data management, data lineage and workflow, and electronic publishing.

This discussion highlights some of the issues currently under consideration in the design of VREs. ClearlyVREs may vary in their technology and collaboration patterns. However, broad similarity among research collaboration issues may demand similar virtual research structures. It is useful to specify types of VREs that might include specific "bundles" of technologies. Such a typology will provide us with a structured view to study, analyze, and compare issues related to communication and collaboration patterns and technology use. As noted previously, there has not been a systematic classification of VREs based on technology. Instead prior typologies tend to emphasis function of the VRE. We now turn to these typologies.

Existing Typologies of VREs

From a functional viewpoint, existing VREs can broadly be classified into three categories (Benford et al., 2001; Finholt & Olson, 1997; Kouzes et al., 1996):

- Distributed research: Virtual environments that facilitate research through the manipulation of physical or biological materials. These are geographically dispersed virtually-connected laboratories where scientists collaborate to conduct research

- Shared instruments: Virtual environments that facilitate observations, modeling, and simulations. These VREs provide instruments for data generation, data collection, and/or data analysis that can be accessed virtually
- Data systems: Virtual environments that facilitate data sharing and data manipulation of already collected data. These are either 'community data storage' that provides a repository of data or 'data analysis infrastructure' that provides community data storage with data analysis tools.

Distributed research often incorporates a data system and sometimes shared instruments in their research environment. Similarly, it is possible to identify shared instrument systems incorporated with data systems.

Finholt and Olson's (1997) collaboratory concept includes distributed and media rich network connections that link people to each other, to facilities, and to information systems. Similar analysis can be found in Kouzes, Myers, and Wulf's (1996) classification of Internet based scientific collaboration. According to them, the Internet based facilities provide (i) a repository of shared or stand-alone data and (ii) access to scientific instruments from distant locations, and (iii) a shared interaction space across several laboratories.

Bos et al. (2007) identified seven types of collaboratories. These are: (i) shared instrument, (ii) community data systems, (iii) open community contribution system, (iv) virtual community of practice, (v) virtual learning community, (vi) distributed research center, and (vii) community infrastructure project. Their classification stands as the most

Table 4 Definitions of Collaboratory

Type of Collaboratory	Definition
Shared Instrument	This type of collaboratory's main function is to increase access to a scientific instrument. Shared Instrument collaboratories often provide remote access to expensive scientific instruments such as telescopes, which are often supplemented with videoconferencing, chat, electronic lab notebooks, or other communications tools.
Community Data Systems	A Community Data System is an information resource that is created, maintained, or improved by a geographically-distributed community. The information resources are semi-public and of wide interest; a small team of people with an online file space of team documents would not be considered a Community Data System. Model organism projects in biology are prototypical Community Data Systems.
Open Community Contribution System	An Open Community Contribution System is an open project that aggregates efforts of many geographically separate individuals toward a common research problem. It differs from a Community Data System in that contributions come in the form of work rather than data. It differs from a Distributed Research Center in that its participant base is more open, often including any member of the general public who wants to contribute.

Table 4 (*Continued)*

Type of Collaboratory	Definition
Virtual Community of Practice	This collaboratory is a network of individuals who share a research area and communicate about it online. Virtual Communities may share news of professional interest, advice, techniques, or pointers to other resources online. Virtual Communities of Practice are different from Distributed Research Centers in that they are not focused on actually undertaking joint projects. The term ‘‘community of practice’’ is taken from Wegner and Lave (1998).
Virtual Learning Community	This type of project’s main goal is to increase the knowledge of participants but not necessarily to conduct original research. This is usually formal education, i.e., provided by a degree-granting institution, but can also be in-service training or professional development.
Distributed Research Center	This collaboratory functions like a university research center but at a distance. It is an attempt to aggregate scientific talent, effort, and resources beyond the level of individual researchers. These centers are unified by a topic area of interest and joint projects in that area. Most of the communication is human-to-human.

Table 4 (*Continued)*

Type of Collaboratory	Definition
Community Infrastructure Project	Community Infrastructure Projects seek to develop infrastructure to further work in a particular domain. By infrastructure we mean common resources that facilitate science, such as software tools, standardized protocols, new types of scientific instruments, and educational methods. Community Infrastructure Projects are often interdisciplinary, bringing together domain scientists from multiple specialties, private sector contractors, funding officers, and computer scientists.
Expert Consultation	Expert consultation provides increased access to an expert or set of experts. The flow of information is mainly one way, rather than two way as in a distributed center.

Note: Quoted from Bos et al. (2007) and Olson (2004).

comprehensive scheme based on functionality. Olson (2003) identified expert consultation as another collaboratory type (see Table 4 for definitions).

These typologies are useful in that they define VREs based on what they do. However, it is difficult to ascertain how technologies map into them. Moreover, based on my search for VREs, it is apparent that the same VRE may fit in more than one of the Bos et al. categories. These current typologies wills serve as useful comparison points for the empirically-derived typology that this study will derive based on technological configurations. We now turn to technologies that might be incorporated into VREs and attempt to develop a conceptual foundation.

Technologies

First, we will consider some issues involved in selecting technologies for VREs. We can argue that a complete VRE should provide scientists with all the technological features necessary for a comprehensive scientific collaboration. We can define comprehensive scientific collaboration in terms of three features: initiation of a discovery process, the scientific discovery experimental processes, and documentation and dissemination processes. These processes are not linear and often overlap.

Initiation of a discovery process phase should emphasize two distinctive processes: ‘social collaboration’ and ‘knowledge processes and idea generation’. Social collaboration includes peer-to-peer communication. Facilitating communication between or among scientists is just a small part of social collaboration. The technology should help scientists to explore the virtual domain to discover other user’s research interests, their scientific products (including primary research materials if available), and provide a way to communicate with them. These tools will also facilitate the development of interpersonal relationships among scientists. In the knowledge processes and idea generation phase, technology should provide means to maximize breadth of document search, access to documents (academic articles, research reports, research data, etc.), research processes, a way to organize and modify all the available materials, and to share all these original and modified materials with others. The technology should also help scientists to store and organize all their communications with their peers. This communication is a part of the discovery process. It will help others to understand the ideas behind the research, phases of idea generation, and also, if needed, will help collaborators to look back into their conversation.

Technology in the scientific discovery processes should provide adequate access to experimental, modeling, visualization, simulation, and data analysis tools. Scientists should be able to use these tools from their remote locations individually and/or collaboratively. There are a growing number of such facilities. However, they have technological limitations and also not all the tools current technology can provide are available.

The documentation and dissemination process starts during the initiation of a discovery process phase. Scientists need to have extensive technologies for documentation. These technologies should enable them to document and edit reports, communications, research materials, and data whenever necessary individually and/or collectively. Technology should also provide facilities to disseminate these documents to peers, supervisors, and to the general public. Technology should also aid gatekeeping and safeguarding these documents if scientists so desire. The documentation tools will not only help scientists to generate and disseminate scholarly materials, but also non-scholarly materials to introduce themselves to the scholarly and the public domain.

What principle of sharing a VRE should use is a major question. The old shared interface system of ‘what you see is what I see’ is no longer applicable in many collaboration settings. When two people need to see different aspects of the same data, visualization needs to reflect the interest and role of the collaborators. However, many graphical systems do not take this into account.

Data management techniques bring other limitations. Available technology can facilitate improved search where a simple scan can retrieve any version of the data. Occasionally we need to examine how certain information was derived from data sources. What types of search systems are widely used in the VREs and why is a matter of investigation. Specific search systems can support data lineage (or data provenance) document retrieval. Data lineage refers to history that explains the process of deriving data from particular data storage. Specific search systems allow users to use keywords and field options (time, author, version, etc.) to search for specific data as well as the history of that data.

Whether electronic publishing should be a feature of current VREs or not is another matter of investigation. There are different publishing and archival techniques not directly related to VRE. When we talk about archived raw data and unpublished materials, allocation of credit is a big issue.

As technology plays a major role in VRE, we need to know the limitations posed by technology. In order to promote virtual research environment, we need to understand what can be done and what cannot be done if we go virtual. It will lead us to understand the degree of research collaboration achievable through the VRE systems.

In the communication literature it is accepted that the presence of a technology and the appropriateness of that technology within the system can heavily influence the effectiveness of collaboration (Poole & Ahmed, in press). In order to investigate the relationship between presence of technology and collaboration effectiveness one must classify VREs based on technology and measure their collaboration effectiveness.

Rationale and Research Questions

The objectives of this examination of VREs are to (1) identify the technology in use within the VRE system, (2) classify VREs based on the configuration of technologies they include, and (3) assess how effective different types of VREs are in promoting collaborative research. Addressing this last task requires us to measure collaboration effectiveness and pattern based on authorship patterns in reports and articles produced in the VRE.

This research will draw its VREs from four scientific domains: physical science, biological science, natural science, and multidisciplinary inquiry. Domain-specific research has a relationship with VRE technology, and therefore, with VRE categorization. Though a significant number of VREs are operating in different parts of the world, we cannot characterize them as most of the VREs have developed based on specific research needs, reached their mature states based on technology usage trial and error, and there are no accepted technology standards. Based on the scientific collaboration and virtual technology literatures, we can argue that scientific collaboration in the virtual environment should have some similarities and some differences across disciplines.

We can see that there are different technologies present in the VRE systems, that collaboration patterns encouraged by the technologies available are different, and that the missions of different research environments also differ. Because of this complex variation, we currently do not have a comprehensive classification for virtual research environments that takes functionality, communication, and collaboration into account. Considering functionality alone tells us only one side of the story, and it is possible that the classification schema will change if we take communication and collaboration technology into account. Based on the preceding discussion, I propose the following research question:

RQ 1: What types of VREs exist based on functionality, communication, and collaboration technologies?

This question will be addressed by empirically deriving a typology of VREs based on the functional, communication, and collaboration technologies they incorporate. This will be done by coding which of a possible list of technological features are present in the VRE and utilizing cluster analysis to identify types.

How technologies actually influence research is one of the key questions in understanding collaboration. Analysis of VRE portals can tell us the nature of technology arrangements within a VRE: how the communication, access to instruments and data, and documentation and document dissemination processes are handled. This leads to a second question:

RQ 2: What are the arrangements of the three types of technologies in VREs?

This question will be addressed by conducting case studies of selected VREs from each type identified in the cluster analysis.

Technology in the scientific discovery experimental processes should provide adequate access to experimental, modeling, visualization, simulation, and data analysis tools. Scientists should be able to use these tools from their remote locations individually and/or collaboratively. There are a growing number of such facilities. However, they have technological limitations and also not all the tools current technology can provide are available. The documentation and dissemination process starts during the initiation of a discovery process phase. Scientists need to have extensive technologies for documentation. These technologies should enable them to document and edit reports, communications, research materials, and data whenever necessarily individually and/or collectively. Technology should also provide facilities to disseminate these documents to peers, supervisors, and to the general public. Technology should also aid gatekeeping and safeguarding these documents if scientists so desire. The documentation tools will not only help scientists to generate and disseminate scholarly materials, but also non-scholarly materials to introduce themselves to the scholarly and the public domain.

We can measure the effectiveness of VRE by looking at the number of publications each VRE is producing. We can also look at the number of collaborators within each virtual environment. The scientific productivity index based on the publication and participation index can tell us their relationship with the technology present within the environment or the classification the fall into. Therefore, we can argue that a comparison between the technological and productive nature of any VRE will tell us about the effectiveness of a particular environment. This will allow us to address a third research question:

RQ 3: What makes collaboration in virtual research environments effective?

CHAPTER III

METHOD

The literature review discussed factors that contribute to the development of a collaborative scientific research environment as identified by Pelz and Andrews (1976). We have also discussed Chin et al. (2002) five identified phases of scientific research group formation and Schur et al.'s (n.d.) identified four types of collaborations that occur during mediated communication. These discussions actually explained different phases of the of the scientific discovery process and provided a basis to evaluate the requirements of technology. We can argue that the effectiveness of the virtual research environment depends upon the presence of technology that supports different phases of the of the scientific discovery process.

These VRE systems include numerous communication and collaboration technologies. The Contextual Resource Evaluation Environment (CREE) feasibility study (Awre & Ingram, 2005) has identified a comprehensive list of available communication and/or collaboration technologies. These are: Announcements, Blogs, Calendar, Chat, Discussion boards, Email, Instant messaging, Interactive learning materials, Place to access and manage content, Place to share files, Polling, Shared bookmarks, Video conferencing, Whiteboard, and Wikis. However, there are other frequently used technologies within VRE systems.

The literature review suggests that the properties of VRE technology will fall under five specific categories areas including several subcategories. These are:

1. Peer-to-Peer Communication/Collaboration Technology: technologies that facilitate dyadic and group communication of individuals of the scientific community (ie, e-mail system, chat-rooms, bulletin boards, etc.)
2. Scientific Instruments: scientific research instruments and tools to generate, collect, manage, and analyze data (ie, space telescope, shared seismographic instruments, etc.)
3. Databases and Data Stores: open, shared, or controlled repository of raw or manipulated data within the VRE system. It does not include generated reports, articles, white papers, or other finalized research documents.
4. Internal Search Systems: systems that facilitate scientific data and document search within the VRE
5. Information Sharing Systems: information sharing aspects will include the following: (i) document dissemination systems (report publishing, information sharing about VRE goals, ongoing research projects, future research projects, scientist profiles, and organizational information) and (ii) links to external resources (instruments, databases, websites, blogs, external search engines, communication tools, and VREs).

Among these five categories, scientific instruments, databases and data stores, search systems, and information sharing systems together will allow any stand-alone user to use scientific facilities to conduct research. However, peer-to-peer communication and collaboration technology allows multiple users to conduct experiments together, as well as to communicate with the VRE. Furthermore, we can distinguish between communication and collaboration technology features. Communication technology features can be defined as features that allow VREs to communicate research, project, and collaborator information to a potential user as well as facilitating communication between the VRE and user. Collaboration technology features can be defined as features that allow multiple users to share research information with each other, run collaborative sessions, and to develop discussion environments. Therefore, for the purpose of the research, the technologies within a VRE will be classified into the general categories of functional, communicative, and collaborative technology.

Based on the review of VREs and the classification method, the study will use the following characteristics to explore VREs:

- Functionality –The attributes of functionality will identify VRE technologies (items) from a functional viewpoint. These items help users to gather information and data, analyze data, develop and run research projects.
- Communication – The attributes of communication will identify VRE technologies that primarily help researcher to communicate with the VRE and its participants. It also includes the intended communication about the VRE research and projects, research participants, and research processes.
- Collaboration - The attributes of collaboration identify VRE technologies that allow multiple users to share research information with each other, run collaborative sessions, and to develop discussion environments.

Sample Selection

In sampling we took the existing VRE dimensions into account. We sampled VREs in each of the four categories within the existing seven fold functional categorization of Bos et al. (2007) which Olson (2004) identified as research-focused VREs: Distributed Research Center (DRC), Shared Instrumentation (SI), Community Data Systems (CDS), and Open Community Contribution System (OCCS).

One challenge in drawing a sample was developing a sampling frame of VREs across different disciplines and geographic regions. The "Collaboratories at a Glance" project offers the only existing comprehensive list of VREs (scienceofcollaboratories.org). The project listed and classified VREs according to Olson et al.'s classification scheme. Within their listing of research-focused VREs 47% fall under Distributive Research Center, 19% are Shared Instruments, 29% are Community Data Systems, and 5% are Open Community Contribution Systems.

While the "Collaboratories at a Glance" project listing was taken as the primary base of information, an extensive web-based search was conducted to identify additional VREs using the following keywords: *VRE*, *collaboratory*, *distributed research environment*, *e-Science*, *virtual research infrastructure*, *research + cyberinfrastructure*, *virtual research*, *e-research*, *online research environment*, *virtual scientific collaboration*, and *online research collaboration*. Links and documents identified by web exploration helped identify existing VREs. About 310 VREs were identified through a preliminary analysis. The percentage distribution identified by the Collaboratories at a Glance project also was supported by the online examination.

A total of 31 VREs (14 DRC, 9 CDS, 6 SI, and 2 OCCS) (10% of the total sampling frame) were randomly selected from the list for online examination. This sampling pattern corresponded to the distribution of VREs found by the Collaboratories at a Glance project.

Collecting Technological Feature Data

The online examination of VRE explored VRE portals to list the technologies they made available. The exploration used the technique of usability inspection. Usability inspection methods are used to understand a website's structure, quality, and usefulness (Nielsen & Mack, 1994). The usability inspection method was selected as the same process comprehensively identifies the features of websites.

Usability inspection is a set of methods that allows an evaluator to inspect a web or a software interface to find design problems affecting usability. This process identifies components of a website, their relationships, and ease-of-use. Usability inspection methods provide us a consistent analysis and evaluation framework to evaluate websites. Olsina, Godoy, Lafuente, and Rossi (1999) developed a specific method of usability inspection to identify and analyze Quality Characteristics and Attributes for Websites. Their aim was to develop a hierarchical and descriptive specification framework for characteristics, sub-characteristics and attributes of academic websites. They mention that the software (or website) quality may be evaluated by usability, functionality, reliability, efficiency, portability, and maintainability characteristics. However, the authors argue that the combination of characteristics needed for the evaluation is based on the study and domain objective and may vary. For their academic website analysis they used a combination of usability, functionality, reliability, and efficiency.

Olsina et al.'s (1999) framework depends upon quality characteristics and attributes that are directly measurable. The primary goal of the framework is to classify and group the elements that might be part of a quantitative evaluation, comparison, and ranking process in a requirement tree. It starts with high-level quality characteristics (i.e., usability, functionality) that provide a conceptual framework of quality requirements. Then these characteristics are decomposed in multiple levels of sub-characteristics, and finally, a sub-characteristic into a set of measurable attributes.

In order to follow this method a hierarchical and descriptive specification framework must be developed. The components of a framework include a combination of the following items: Title, Code, Type, Higher level Characteristic, Sub-characteristics, Attribute, Definition / Comments, Model to determine the Global/Partial Computation, Employed Tool/s, Preference Scale, and Examples.

As the development of the framework is dependent upon the scenario, based on the VRE discussions we can argue that an adoption of this method to analyze VRE should take the following components in developing a framework: Higher level Characteristic, Sub-characteristics, Attribute, and Definition (see Table 5 for complete framework). Discussion about Preference Scale is not necessary as our elementary coding criterion is binary asking only for data about availability (presence 1; absence 0) of technology. The data collection type is manual and observational.

Observations of VRE portals allowed us to list features available within a system. First, this research started identifying and listing the major types of features incorporated within a system. Second, each of the identified features was defined based on their characteristics. Third, the research kept on expanding the list and defining newly identified features. After a certain time, we could no longer find new features. The exhaustive list was used as the final feature list for data collection. Fourthly, sub-categories were developed and defined based on the exhaustive list. Finally, sub-categories were placed under already defined categories. A total of 25 features were identified through this exploration. These features were combined into 11 subcategories.

The first set of observed features was related to VRE functionality. VRE provides a distinctive portal as it includes features that aid scientific enquiry. Databases, scientific instruments, and scientific literature depository allow users to conduct scientific experiment or observation, collect and analyze scientific data, and to explore existing data and literature. Scientific instruments help conduct experiment, observe phenomenon, collect data, and to analyze data. Technology that aids experiment and observation for data collection were labeled here as scientific instruments. Besides, these instruments there are online and offline data analysis and visualization tools. Online tools are embedded within the VRE grid. VREs also provide software downloadable to a computer for the same purpose. In most cases, functional ability of downloadable software is lower than the online software.

VREs often include data and literature depositories. Data depositories are collection of already collected data from any ongoing or completed project. Biological sciences also deposit models in their data depository. Literatures include scientific literatures, VRE related documentations, technical reports, and instructions to use instruments and data.

The majority of the VREs include search options within their portal. Search options allow users to search for instruments, data, literature, or individuals affiliated with research. Internal search option allows conducting search only within the site. External options allows to search outside the site. Often, external search option embeds a Google search within the site to allow users to conduct an external search without leaving the site.

Links connect different part of a portal together. Links may be internal by only tying resources of a VRE together or external by allowing users to leave the site and browse external resources. External links are often directed either to another VRE or to tools and datasets of external VRE. In most cases, VRE provide links to sister projects, VREs with similar objectives, VREs in the same field, and related datasets. Sometimes, VREs share resources like software, analysis tools, or instruments among them. In those cases, external links connect users to those external resources. Sometimes, VREs do not provide all the resources need for scientific discovery. In that case, link to related resources can enhance the discovery process. An example could be a data depository VRE without any analysis tool providing link to an analysis tool of another VRE. External links, therefore, develop a network of resources.

Table 5 Classification Scheme

Category	Sub-category	Attribute	Definition
Functionality			VRE technologies (items) from a functional viewpoint. These items help users to gather information and data, analyze data, develop and run research project.
	Links		Hyperlinks that allow user to explore information, data, tools, and communicative and collaborative features.
		Internal links	Hyperlinks within the VRE site.
		External Links	Hyperlinks outside the VRE site.
		Link to external database	Hyperlinks to a database that is not a part of the VRE project.
		Link to external tools	Hyperlinks to a tool that is not a part of the VRE project.
		Link to external VREs	Hyperlinks to another VRE project information; not necessarily to databases or tools.

Table 5 (*Continued*)

Category	Sub-category	Attribute	Definition
		Significant number of missing links	Any VRE with 4 or more missing links.
	Search		Options to look for information, data, or links.
		Search option global	One field search option that uses keyword input to look for information.
		Search option specific	Multiple field search option that allows user to use a combination of multiple keyword and already set fields.
		Search option external	Option that allows user to search outside the VRE portal.
	Depository		Collection of scientific documents and data.
		Database	Collection of data files.
		Literature	Collection of scientific reports and publications.

Table 5 (*Continued*)

Category	Sub-category	Attribute	Definition
	Instruments		
		Data analysis tools	Online software that helps data analysis and visualization.
		Downloadable software	Downloadable (offline) software that a person can download to a computer and use for data analysis or visualization.
		Scientific Instruments	Accessible scientific instruments for data collection, material analysis, or observation.
Communication			VRE technologies that primarily help researcher to communicate with the VRE and its participants. It also includes the intended communication about the VRE research and projects, research participants, and research processes.

Table 5 (*Continued*)

Category	Sub-category	Attribute	Definition
	Site	Site information	Information about the VRE portal including mission statement, objectives, creation, maintenance, and available facilities.
	Research	Project information	Information about ongoing, completed, or future research projects of the VRE.
		Research information	Summary of objectives, mission statement, present condition, findings, and future directions of a specific research project.
	Help	FAQ	FAQ help for using VRE, data, user access, instruments, communication, or collaboration.
	Collaborators	Collaborator information	List of individual or institutional participants
	E-mail	Global e-mail	General e-mail option to VRE
		Specific e-mail	E-mail option to specific people related to management, coordination, and research.

Table 5 (*Continued*)

Category	Sub-category	Attribute	Definition
Collaboration			VRE technologies that allow multiple users to share research information with each other, run collaborative sessions, and to develop discussion environments
	Asynchronous		Non real-time collaboration between scientists.
		Blog	Blog
		Wiki	Research based wiki
		Newsgroup	Newsgroup option for participants.
	Synchronous		Real time collaboration option.
		Chat	Chat room option.
		Conferencing	Videoconferencing or teleconferencing options.

VREs provide information related to the side development, objectives, mission statement, research information, information about completed, ongoing, and future projects, and participating individuals and institutions. The information helps a user to know about the site, its research projects, and resources offered by the VRE. Sometimes the VRE includes a FAQ section. Furthermore, a user can communicate to the VRE for more information or to get access to the restricted resources using the e-mail feature provided in the site. Almost all the VREs have a global e-mail functionality that allows users to communicate with a designated management/coordination authority. Often, VREs provide e-mail to specific persons – management/coordination personnel, affiliated scientists, team leaders, project managers, or technologists. These features communicate essential information to VRE users and provide means of further communication if required by the user.

VREs also provide synchronous and asynchronous collaboration tools. The online exploration identified blogs, wikis, chat rooms, newsgroups, and teleconferencing systems in different VREs. Collaboration tools allow a user to join scholarly discussion and to be a part of a networked scientific community. Blogs and newsgroups allow users to post an idea or opinion and join or observe an ongoing discussion. Wikis allow users to post an entry about any scientific topic, idea, definition, or explanation. Blogs, newsgroups, and wikis, therefore, aid the development of scientific collaboration network. They also provide a collaborative social environment. Chat rooms and conferencing systems aid real time discussion among collaborators. Depending upon the requirements, any dispersed scientific group can have a text only (i.e., chat room), voice only (teleconferencing), or audio-visual (videoconferencing) discussion. Using the appropriate technology, scientists can present their idea or result to their group or community and get real time feedback. Conferencing technology can also help arrange lecture sessions crucial for knowledge transfer within the scientific community.

Collecting Productivity Data

This research defines productivity using three different components: the amount of publications produced by a particular VRE, the number of collaborators engaged in producing those publications, and collaborators affiliations. The amount of collaboration would allow us to observe the productive nature of a VRE. Another, major idea behind this analysis was to explore the level of collaboration by identifying the number of collaborators as well as identifying the number of internal and external collaborators. The number of internal collaborators would allow us to understand if the VRE is allowing scientists to form groups to work together. External collaborator information would allow us to observe the nature of the scientific group – if the network is expanding outside the VRE participants.

The nature of the technological feature data collection method provided a simple way to collect productivity data. The presence of a literature depository allows us to ascertain the presence of publications within a VRE. Feature data collection also tells about the presence of collaborator information within a VRE. While collecting feature data, both types of data will be collected in depth. The number of publications and collaborators were collected and used for productivity analysis. First, the total numbers of publications were collected. Journal articles, books, scientific reports, and conference presentations were considered as publications. Other types of identified publications such as minutes, technical documentations (i.e., how to operate an instrument), VRE missions and objectives were not considered as scientific publications. Second, 25 sample publications from 2005 to 2009 were selected from VREs having with both collaborator and publication information. 25 publications gave us a benchmark as the majority of portals had equal to or less than this amount of publications. The number was also enough to include majority of the authors participating within a VRE. All the publication of a VRE were included if it did not have 25 publications. A VRE without either publication or collaborator information were not considered for this analysis. Second, a person was considered as an internal participant if his/her name was listed as a participant or his/her affiliated institution was listed as a participating institution. When an authors name was not listed in the VRE and the institutional affiliation data was not provided, a Google search was employed to indentify author's institutional affiliation.

Data Analysis

Cluster analysis using Ward's method was employed to categorize VREs. The aim of the cluster analysis was to investigate if we can divide VREs into meaningful subgroups based on their functionality, communication, and collaboration features.

Wards method is a hierarchical clustering that minimizes the loss of information during merging. Minimizing loss of information was particularly important because of the small sample size. "At each step of ward's method, the union of every possible pair of groups is considered and two groups whose fusion results in the minimum increase in the loss of information are merged" (Gan, Ma, & Wu, 2007, p. 133). Gan et al. (2007)

state that Ward's method usually produces very good approximation if we use squired Euclidian distance to compute the dissimilarity matrix. This analysis used Ward's method with squired Euclidian distance for binary data.

This research also used Pearson's correlation to investigate relationship between features, features and productivity items, and feature subgroups and productivity items.

CHAPTER IV

RESULTS

This chapter consists of three sections. The first section describes general characteristics of the sample and the cluster analysis. The second section provides descriptive analysis of each cluster and case studies of two extreme cases within the cluster. The final section reports results on the productivity of each VRE in the various clusters and as it relates to comprehensiveness of the tools in the VREs.

Cluster Analysis

A total of 31 VREs were sampled based on the Olson et al' classification. The sample consists of 13 distributive research centers, 6 shared instruments, 10 community data systems, and 2 open community contribution systems. From the domain perspective majority of the VREs, fifteen, are Biological/Biomedical Sciences followed by eight from the Natural Sciences, four from the Physical Sciences, and four from Multidisciplinary Sciences. The sample of the Biological/Biomedical VREs does not include any open community contribution system VRE. Both the open community contribution system VREs are in the Natural Science domain. Physical Science and Multidisciplinary Science VREs are either distributive research center or shared instrument types. The Natural Science domain is the only one that includes all four types of VREs (see Table 6).

We subjected the sample to a hierarchical cluster analysis using Ward's procedure and this yielded five clusters using a cut point at 6. Cut point at 6 provides us with a manageable amount of clusters where the distances between clusters are also significant. Before that point, smaller several smaller clusters were very close in characteristics. After that point, most of the VREs merged into a couple of clusters following the procedures of Ward's method.

Table 6 Type and Domain of VRE by Cluster

Cluster	VRE Type				VRE Domain			
	DRC	SI	CDS	OCCS	BIO	PH	NAT	MULTI
1	5	1	2	0	2	2	3	1
2	0	2	2	0	3	0	1	0
3	4	0	4	1	6	1	2	0
4	1	1	1	1	3	0	1	0
5	4	2	0	0	1	1	1	3
Total	14	6	9	2	15	4	8	4

Note: DRC: Distributed Research Center; CDS: Community Data Systems; SI: Shared Instrument; OCCS: Open Community Contribution System; BIO: Biological/Biomedical; NAT: Natural; MULTI: Multidisciplinary; PH: Physical.

Cluster 1 consists of eight VREs, three from the field of Natural Sciences, two each from Physical and Biological/Biomedical Sciences, and one from the Multidisciplinary Sciences. Half of those VREs are Distributed Research Centers, three are Community Data Systems, and one is a Shared Instrument. Cluster 2 consists of four VREs, three from the field of Biological/Biomedical Sciences and one from Natural

Science. Half of those VREs are Shared Instruments and the other half are Community Data Systems. Cluster 3 is the largest among clusters with nine VREs, six from the Biological/Biomedical Sciences, two from the Natural Sciences, and one from the Physical Sciences. Four of those VREs are Distributed Research Centers, four Community Data Systems and one is an Open Community Contribution System. Cluster 4 consists of four VREs, three from the Biological Sciences and one from the Natural Sciences. Interestingly, this cluster has one representation of all four VRE types. Cluster 5 consists of six VREs, three from the Multidisciplinary Sciences and one each from the Natural, Physical, and Biological/Biomedical Sciences. Four of the VREs are Distributed Research Centers and the remaining two are Shared Instruments.

There is a lot of variation of features within these clusters. Cluster 3 was the most comprehensive VREs. Cluster 1 is the second most comprehensive and the only cluster where all the VREs have all three components. Cluster 2 was identified as functional with communication option, but lacking collaborative options. In fact, the functional and communicative comprehensiveness of cluster 2 is higher than cluster 1. Cluster 4 and cluster 5 are low in functionality and communicative aspects and do not have any collaboration features. Cluster 5 is identified as the least developed cluster. Comprehensiveness was not related to VRE types or fields.

In terms of their general constellations of characteristics, we called Cluster 1 the Full-Service Balanced, Cluster 2 Simple Non-collaborative, Cluster 3 Full-service Functional, Cluster 4 Low-function Non-collaborative, and Cluster 5 Non-functional Informative. Case studies of clusters were conducted to explore the nature of the clusters more fully. The case studies are based on the variations of functionality, communication, and collaboration features, and the description of two extreme cases: most comprehensive and least comprehensive VREs in each cluster.

Descriptive Analysis of Clusters

Cluster 1: Full-Service Balanced

Cluster 1 provides a set of VRE with moderate comprehensiveness. Though there are differences among the VREs within this cluster, what grouped them together is the presence of at least one common feature of the three surveyed areas (functionality, communication, and collaboration). Hence, this cluster is balanced in providing full services for users. The breakdown of the VREs from the most to the least comprehensive (based on the number of features present) is presented in table 7.

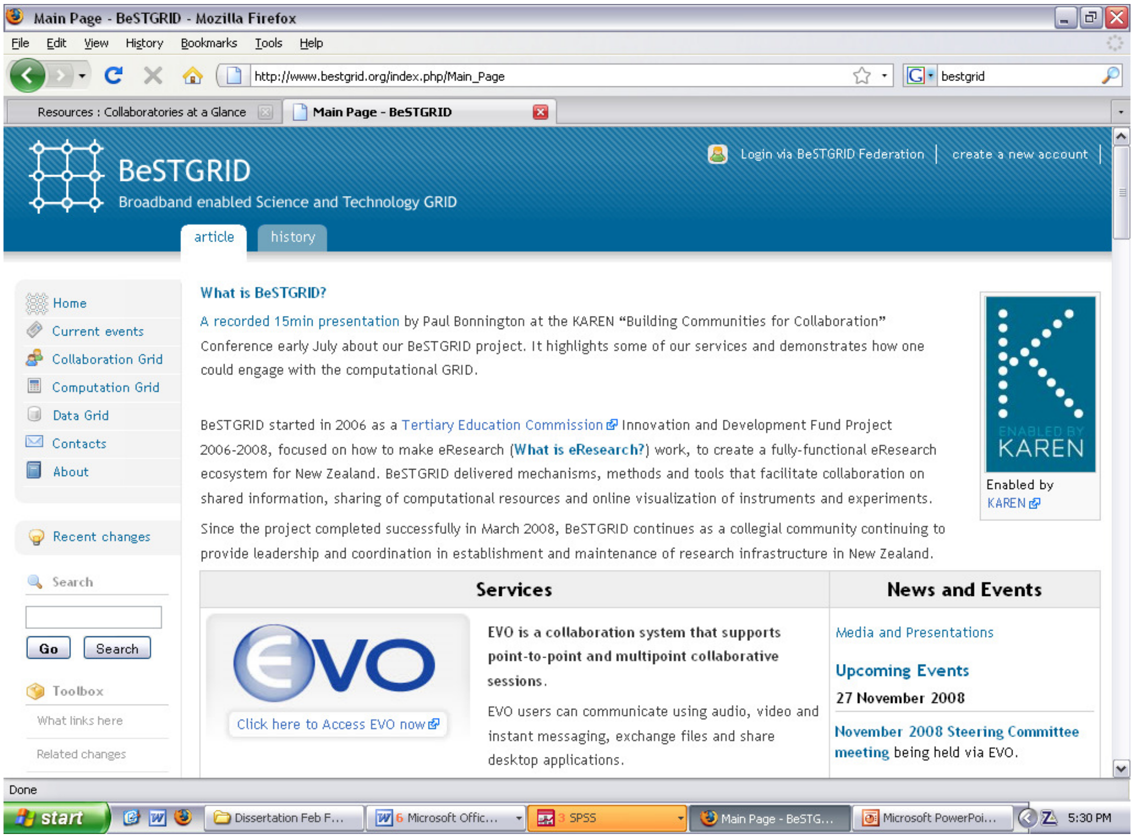
BeSTGRID came out as the most comprehensive VRE by having 19 of the 25 possible features combining functionality, communication, and collaboration, whereas, The Virtual Environments for Research in Archaeology (VERA) provides the least. This section will present case analyses of these two “extreme” members of cluster 1.

Table 7 Breakdown of Cluster 1 VREs from the Most to the Least Comprehensive

VRE	Total Features
BeSTGRID	19
CHRONOS	16
Biological Collaborative Research Environment (BioCoRE)	16
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	16
Theoretical and Computational Biophysics Group (TCBG)	16
International Virtual Observatory Alliance (IVOA)	15
National Fusion Collaboratory	14
The Virtual Environments for Research in Archaeology (VERA)	13

BeSTGRID

New Zealand’s Tertiary Education Commission Innovation and Development Fund Project started the development of BeSTGRID (Broadband enabled Science and Technology GRID) eResearch system in 2006 (Figure 2). The focus of the project was to develop a fully-functional eResearch ecosystem that would facilitate sharing of computational resources through shared information, tools, and online visualization instruments. The initial project was completed successfully in 2008. Since then, BeSTGRID has provided an eResearch infrastructure for the education and research community in New Zealand.



What is BeSTGRID?

A recorded 15min presentation by Paul Bonnington at the KAREN "Building Communities for Collaboration" Conference early July about our BeSTGRID project. It highlights some of our services and demonstrates how one could engage with the computational GRID.

BeSTGRID started in 2006 as a Tertiary Education Commission Innovation and Development Fund Project 2006-2008, focused on how to make eResearch (**What is eResearch?**) work, to create a fully-functional eResearch ecosystem for New Zealand. BeSTGRID delivered mechanisms, methods and tools that facilitate collaboration on shared information, sharing of computational resources and online visualization of instruments and experiments.

Since the project completed successfully in March 2008, BeSTGRID continues as a collegial community continuing to provide leadership and coordination in establishment and maintenance of research infrastructure in New Zealand.

EVO is a collaboration system that supports point-to-point and multipoint collaborative sessions.

EVO users can communicate using audio, video and instant messaging, exchange files and share desktop applications.

Media and Presentations

Upcoming Events

27 November 2008

November 2008 Steering Committee meeting being held via EVO.

Figure 2 BeSTGRID portal homepage.

University of Auckland, University of Canterbury, and Massey University are the primary institutions participating in the BeSTGRID system along with several other research and educational institutes. The collaboration environment within the system is divided into three main grids: collaboration, computation, and data.

The collaboration grid is a compilation of three different systems: EVO, Sakai, and Access Grid. EVO is a collaboration system that allows users to use audio, video, instant messaging, file exchange system, and desktop applications sharing systems in point-to-point and multipoint collaborative sessions. Sakai is a set of software tools that provides a set of built-in features to design collaboration. With the help of Sakai, collaborators can develop standard tailored websites to present and exchange information with their peers. Sakai can aid creating tailored bulletin boards, announcement systems, discussion threads, and information and file submission systems within a web environment. Access Grid is a collection of resources that help facilitate collaboration when large format graphic displays, high-level visualization, and interactive visualization environments are needed during collaboration.

BeSTGRID's Computational GRID allows for the sharing of high-performance computational resources. This grid linked supercomputing and mass storage resources located at the supercomputing centers of the University of Auckland, University of Canterbury, University of Otago, and Massey University.

The Data GRID system provides a sharing and management environment for large amounts of distributed data often combined with or linked to the computational grid computing systems. The Data GRID provides collaborators with over 100 Terabytes of research data stored primarily at Auckland, Canterbury and Massey locations.

There is a neat arrangement of the GRIDs and other resources within the VRE portal. The main-page provides a short description of the portal and the VRE, GRIDs, and other technical features accessible through the portal. The main-page links provide access to home, current events, collaboration grid, computation grid, data grid, contacts, about BeSTGRID, and recent changes within the system. The contact link provides e-mail addresses of specific people responsible for projects and management and lists of contact persons for the participating institutions. There are additional links to a file uploading system and a special pages link that provides an alphabetic listing of articles and other documents, topic explanation, key terms used within the system, and technical specifications. There are also links to privacy policy and disclaimer statements. The right hand side of the portal main-page provides a column of news and events with links within the short description for further readings.

The collaboration, computation, and data main links allow visitors to look at specific items and services in that particular category. A list of articles, documents, and help-guides related to the category are provided for all three main GRID links. There are history tabs for all the link pages showing latest upgrades and changes within the system components. This information is particularly useful for user to see the changes within the environment, especially, how the VRE is developing over time.

VERA

Like BeSTGRID, the Virtual Environments for Research in Archaeology (VERA) also has a similar neat arrangement (Figure 3). The portal main-page provides 5 specific tabs: about VERA, events, portal, Wiki, and Blog.

The about VERA tab provides links to welcome notes, contact, get involved, standards, software, people, publications, and links. The content of the welcome link – “Welcome to the VERA Project” is the default main display of the portal. The content of the welcome notes provides further links to the participating institutions, projects, announcements, and blogs. The contact link provides organizational contact information (physical address, telephone, and e-mail) as well as e-mail addresses of specific discussion groups. The standards page provides information about the technical standards adopted by the project for developing the portal, webpages, wikis, and blogs as well as the standard specifications of archeology used by the VERA project.

Welcome to the VERA Project

The **V**irtual **E**nvironments for **R**esearch in **A**rchaeology (VERA) project aims to produce a fully-fledged virtual research environment for the archaeological community. It will address user needs, enhancing the means of efficiently documenting archaeological excavation and its associated finds, and create a suitable Web portal that provides enhanced tools for the user community. VERA aims to develop utilities that help encapsulate the working practices of current research archaeologists unfamiliar with virtual research environments.

The VERA project is being undertaken in the School of Systems Engineering, and the Department of Archaeology, University of Reading, with user testing and analysis being undertaking in the School of Library, Archive and Information Studies, University College London. VERA is funded by the JISC as part of phase 2 of the Virtual Research Environments programme, and runs from April 1st 2007 until March 31st 2009.

The VERA project builds on the successful JISC VRE 1 project,

Figure 3 VERA portal homepage.

The people link provides the names of key people participating in the projects and their affiliations. However, it does not provide any contact information. The publications link provides a list of papers, reports, articles, and presentations. The majority of the mentioned documents are accessible (as .pdf files and slideshows) through the system. The link button provides links to project websites, databases, URLs of partner institutions, and sister projects.

The event tab lists past and forthcoming events and workshop information. The subsections are further hyperlinked to specific descriptions. The portal tab provides a link to the work-in-progress version of the VERA portal. The Wiki page provides restricted (password protected) access to the Wiki system and the Blog tab shows the VERA blog.

Although neatly arranged, the Virtual Environments for Research in Archaeology (VERA) ranked at the bottom of comprehensiveness within the cluster because it is lacking in functionality and communication. The VRE environment does not provide any analysis tool, specific or global search options, or external links to other VREs, databases, or tools. Though there are names of collaborators, information about them (including specific e-mail addresses) are absent in the system. The system does not also provide any options for chat or video-enabled communication and collaboration. Help-files or FAQs are also absent within the system. However, VERA scored higher than other VREs beside BeSTGRID in collaboration features as it provides a Wiki, a Blog, and a newsgroup. Interestingly BeSTGRID, the other highest ranked VRE, does not provide Wiki and Blog options. Instead, BeSTGRID supports collaboration through chat-room and teleconferencing/videoconferencing options.

A breakdown from a functionality perspective reveals five different levels within the cluster (Table 8). BeSTGRID leads the ranking in terms of functionality, having 9 out of 13 surveyed features. The breakdown from a functionality perspective is as follows:

Table 8 Breakdown of Cluster 1 VREs from Functionality Perspective

VRE	Total Features
BeSTGRID	9
CHRONOS	8
Biological Collaborative Research Environment (BioCoRE)	8
Theoretical and Computational Biophysics Group (TCBG)	8
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	7
International Virtual Observatory Alliance (IVOA)	6
National Fusion Collaboratory	5
The Virtual Environments for Research in Archaeology (VERA)	5

The breakdown from a communication perspective creates four different groups with BeSTGRID having all the communication features surveyed. The ranking base on communication features is provided in Table 9.

Although the presence of functionality and communication features within the VREs constituted a similar ranking pattern, collaboration revealed somewhat different ranking. The Virtual Environments for Research in Archaeology (VERA) that ranked at the bottom of both functionality and communication surprisingly tops the collaboration raking along with BeSTGRID by having three out of five collaboration features.

Biological Collaborative Research Environment (BioCoRE) ranks last with only one collaboration feature present within the system. The ranking is provided in table 10.

Table 9 Breakdown of Cluster 1 VREs from Communication Perspective

VRE	Total Features
BeSTGRID	7
National Fusion Collaboratory	7
Biological Collaborative Research Environment (BioCoRE)	7
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	7
International Virtual Observatory Alliance (IVOA)	7
CHRONOS	6
Theoretical and Computational Biophysics Group (TCBG)	6
The Virtual Environments for Research in Archaeology (VERA)	4

Newsgroup is the most prominent collaboration feature within this group of VREs. Only Biological Collaborative Research Environment (BioCoRE) does not have a newsgroup feature within its virtual environment. Newsgroup is followed by Wiki and chat-room features (present within three out of eight VREs).

Among the communication features, all of the VREs have site information, project information, research information, and global e-mail. VERA stands as the only VRE without a specific e-mail feature.

Table 10 Breakdown of Cluster 1 VREs from Collaboration Perspective

VRE	Total Features
BeSTGRID	3
The Virtual Environments for Research in Archaeology (VERA)	3
CHRONOS	2
National Fusion Collaboratory	2
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	2
International Virtual Observatory Alliance (IVOA)	2
Theoretical and Computational Biophysics Group (TCBG)	2
Biological Collaborative Research Environment (BioCoRE)	1

The arrangement of portal, information display, and technological features of other VREs of this cluster falls between BeSTGRID and VERA without much discrepancy. Among those, International Virtual Observatory Alliance (IVOA) possesses all of the communication features but is lacking in terms of functionality and collaboration. Biological Collaborative Research Environment (BioCoRE) stands as the only VRE with just one collaborative function – a chat-room.

Cluster 2: Simple Non-Collaborative

Cluster 2 contains a small set of VREs with moderate functional, weak communicative, and no collaborative aspects. These VREs are simple in nature, allows user to perform some research task, and does not provide enough option to start a collaborative network. There are variations among the VREs within this cluster. However, there is more similarity than differences among VREs in functionality and communication features.

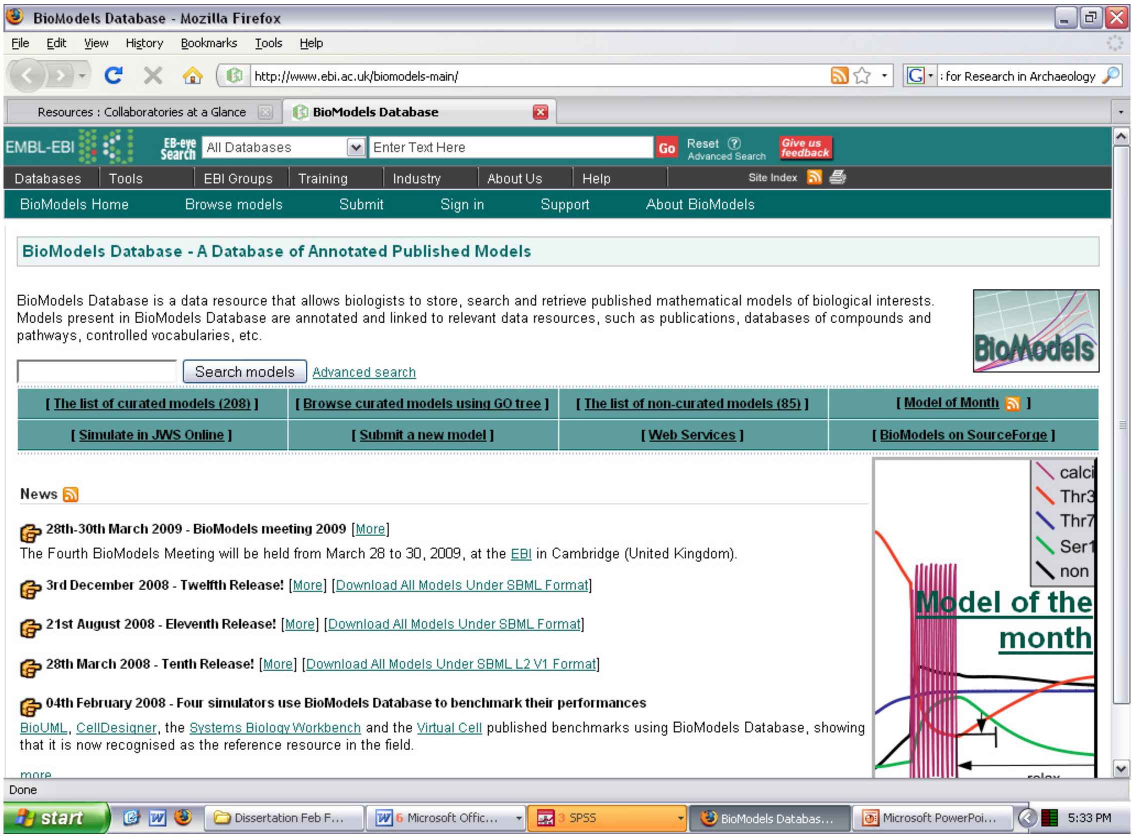
Biomodels Database came out as the most comprehensive VRE within cluster 2 by having 18 of the 25 possible features combining functionality, communication, and collaboration. However, this combination is actually based on functionality and communication features. None of the VREs in Cluster 2 have any collaboration features. The differences among VREs within this cluster are nominal. The breakdown of the VREs from the most to the least comprehensiveness (based on the number of features present) is as follows (Table 11):

Table 11 Breakdown of Cluster 2 VREs based on Comprehensiveness

VRE	Total Features
BioModels Database	18
Sloan Digital Sky Survey (SDSS)	17
Molecular Interactive Collaborative Environment (MICE)	16
Virtual Cell Portal	13

Biomodels Database

Biomodels Database provides a research environment that enables scholars to share mathematical models of biological systems. Scientists can store, search, and retrieve models using the system. The homepage of the database portal provides links to database, tools, training, about the database, help, model browsing and submission of new models (Figure 4). There is a section in the homepage showing briefs of current news with links to more detailed description.



BioModels Database - A Database of Annotated Published Models

BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests. Models present in BioModels Database are annotated and linked to relevant data resources, such as publications, databases of compounds and pathways, controlled vocabularies, etc.

News

28th-30th March 2009 - BioModels meeting 2009 [More]
The Fourth BioModels Meeting will be held from March 28 to 30, 2009, at the EBI in Cambridge (United Kingdom).

3rd December 2008 - Twelfth Release! [More] [Download All Models Under SBML Format]

21st August 2008 - Eleventh Release! [More] [Download All Models Under SBML Format]

28th March 2008 - Tenth Release! [More] [Download All Models Under SBML L2 V1 Format]

04th February 2008 - Four simulators use BioModels Database to benchmark their performances
BioUML, CellDesigner, the Systems Biology Workbench and the Virtual Cell published benchmarks using BioModels Database, showing that it is now recognised as the reference resource in the field.

Figure 4 Biomodels Database portal homepage.

The portal also allows global and specific model search options. Both the global and specific search options allow users to use keywords to search; however, the specific search option allows retrieval of models. The support option provides a FAQ section with answers related to the virtual environment, VRE mission, and biological models.

The about option links users to news, meeting information, information about internal and external contributors, and global contact information.

The tools option provides links to databases, analysis services, and a data integration system for large scale data querying. The database option provides systematic links to browse various internal and external databases. This virtual environment provides excellent database access. However, users do not have access to any scientific instruments. Users of this site can use the previously collected data. However, collection of new data is not possible through this website.

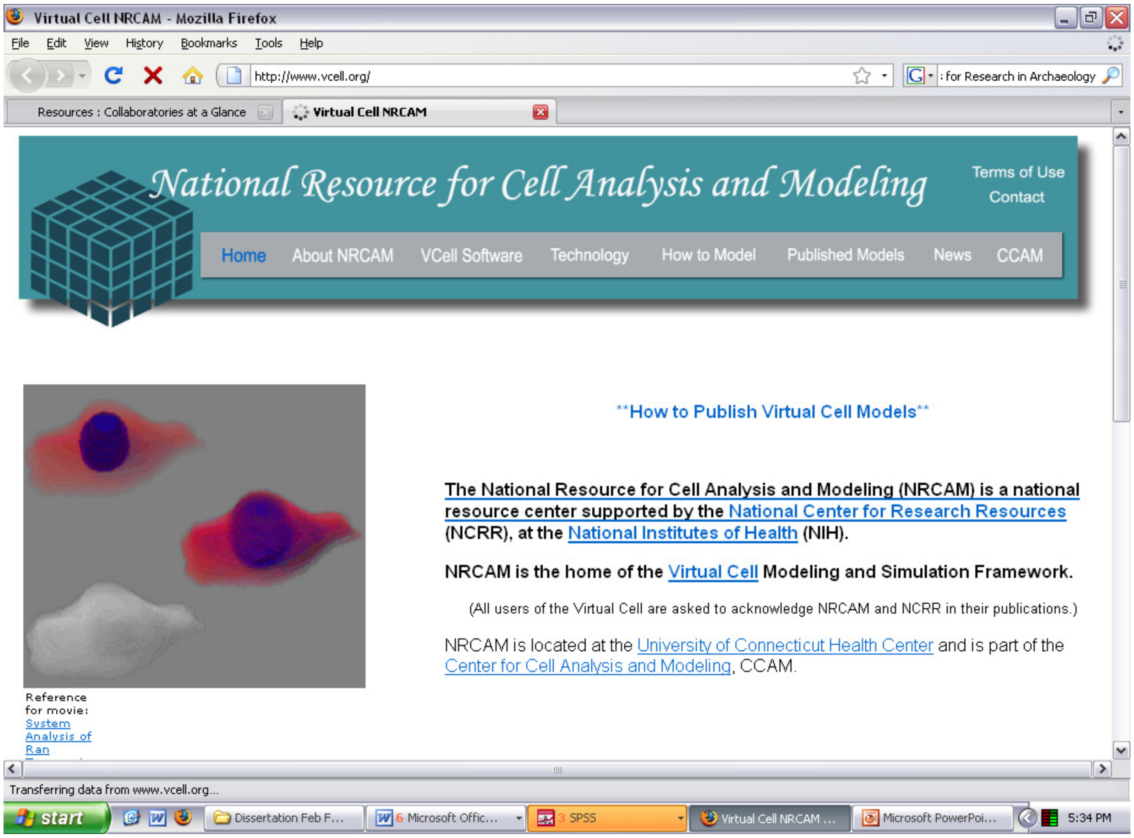
The Virtual Cell

The Virtual Cell is a software modeling environment developed by the National Resource for Cell Analysis and Modeling (NRCAM). It provides a virtual environment for quantitative biological cell research. The objective of this environment is to provide facilities for experimental manipulation and computational simulation ranging from molecular motors to tissue-wide process.

The portal main page provides links to virtual cell software and related software projects, about the environment, technology, how to model instructions, published models, and news (Figure 5). It also provides link to the main NRCAM website.

The technology link lets users access the modeling software. It also provides further links to explanations of technology in use within the environment, user-interface-related information, and descriptions of mathematical and modeling frameworks. The news section provides information about the field related courses and annual meetings.

About NRCAM links user to researchers and administrators of the Virtual Cell including their e-mail and telephone numbers.



How to Publish Virtual Cell Models

The National Resource for Cell Analysis and Modeling (NRCAM) is a national resource center supported by the National Center for Research Resources (NCRR), at the National Institutes of Health (NIH).

NRCAM is the home of the Virtual Cell Modeling and Simulation Framework.

(All users of the Virtual Cell are asked to acknowledge NRCAM and NCRR in their publications.)

NRCAM is located at the University of Connecticut Health Center and is part of the Center for Cell Analysis and Modeling, CCAM.

Figure 5 Virtual Cell portal homepage.

A breakdown from a functionality perspective reveals three different groups within the cluster, with very little differences within levels. The breakdown from a functionality perspective is as follows (Table 12):

Table 12 Breakdown of Cluster 2 VREs based on Functionality

VRE	Total Features
BioModels Database	11
Molecular Interactive Collaborative Environment (MICE)	10
Sloan Digital Sky Survey (SDSS)	10
Virtual Cell Portal	9

The breakdown from a communication perspective also creates three different levels. There is a minimum difference between level 1 and 2. However, the Virtual Cell Portal stands alone with a very low level of communication features, making it the least effective VRE within the cluster in terms of the facilities it provides. The ranking based on communication features is as follows (Table 13):

Table 13 Breakdown of Cluster 2 VREs based on Communication

VRE	Total Features
BioModels Database	7
Sloan Digital Sky Survey (SDSS)	7
Molecular Interactive Collaborative Environment (MICE)	6
Virtual Cell Portal	4

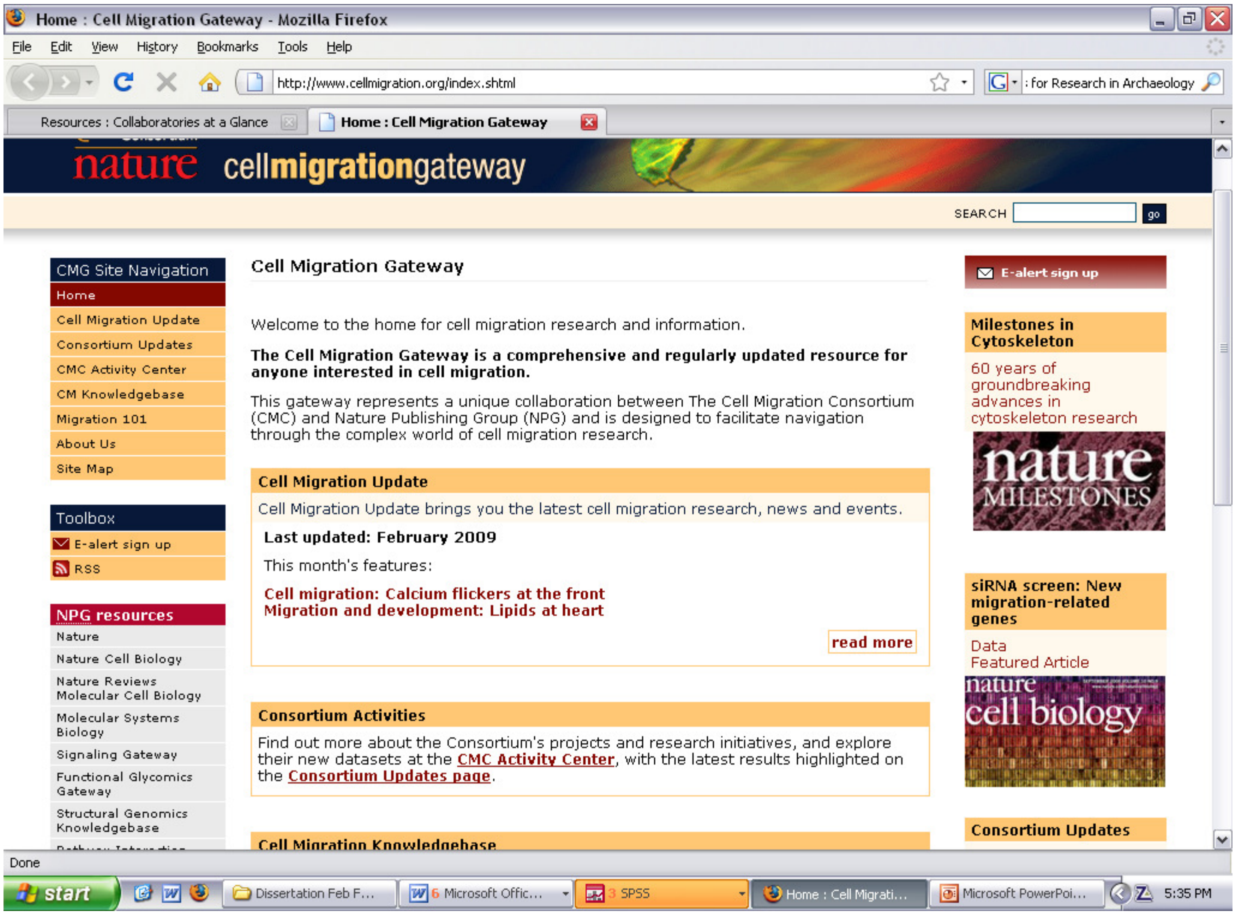
With the exception of the BioModels Database, the functionality of Cluster 3 lacks external links and scientific instruments. Virtual Cell and SDSS provide no links to external databases, VREs, or tools. MICE provides no specific search options. The communication aspect of this cluster is also problematic. Virtual Cell provides no collaborator information. Virtual Cell and MICE provide no FAQ options. Although, the presence of functionality and communication features within the VREs constituted a similar ranking pattern and the differences between two consecutive levels are low, collaboration seems to be the biggest concern for this cluster. Having no collaborative feature within the environment made these VREs less user-friendly.

Cluster 3: Full-Service Functional

Cluster 3, like cluster 1, provides full services to the VRE users. However, unlike the moderate nature of cluster 1, this set of VREs provides high functionality. Cell Migration Consortium came out as the most comprehensive VRE within this cluster by having 23 of the 25 possible features combining functionality, communication, and collaboration. Indeed, Cell Migration Consortium came out as the most comprehensive VRE within all sampled VREs. The second most comprehensive VRE within the cluster, Biomedical Informatics Research Network: Coordination Center (BIRN CC) also came out as the second most comprehensive VRE within all sampled VREs. In fact, cluster 3 has half of the top 10 most comprehensive VREs within the sample. From a comprehensiveness perspective, though all the VREs are somewhat comprehensive, there is a big difference among the VREs in this cluster. In the comprehensiveness ranking, the last three VREs do not have any collaboration features. The breakdown of the VREs from the most to the least comprehensive (based on the number of features present) is provided in Table 14.

Nature--Cell Migration Consortium

Cell Migration Consortium came out as the most comprehensive VRE within all sampled VREs (Figure 6). The Cell Migration Consortium provides a migration-related interdisciplinary collaborative environment for more than 20 participating institutions.



CMG Site Navigation
- Home
- Cell Migration Update
- Consortium Updates
- CMC Activity Center
- CM Knowledgebase
- Migration 101
- About Us
- Site Map

Toolbox
- E-alert sign up
- RSS

NPG resources
- Nature
- Nature Cell Biology
- Nature Reviews Molecular Cell Biology
- Molecular Systems Biology
- Signaling Gateway
- Functional Glycomics Gateway
- Structural Genomics Knowledgebase

Cell Migration Gateway

Welcome to the home for cell migration research and information.

The Cell Migration Gateway is a comprehensive and regularly updated resource for anyone interested in cell migration.

This gateway represents a unique collaboration between The Cell Migration Consortium (CMC) and Nature Publishing Group (NPG) and is designed to facilitate navigation through the complex world of cell migration research.

Cell Migration Update

Cell Migration Update brings you the latest cell migration research, news and events.

Last updated: February 2009

This month's features:

Cell migration: Calcium flickers at the front
Migration and development: Lipids at heart

read more

Consortium Activities

Find out more about the Consortium's projects and research initiatives, and explore their new datasets at the **CMC Activity Center**, with the latest results highlighted on the **Consortium Updates page**.

Cell Migration Knowledgebase

E-alert sign up

Milestones in Cytoskeleton

60 years of groundbreaking advances in cytoskeleton research

nature MILESTONES

siRNA screen: New migration-related genes

Data
Featured Article

nature cell biology

Consortium Updates

Figure 6 Cell Migration Consortium portal homepage.

The Consortium primarily supports workshops on migration-related issues and makes consortium data, reagents, and protocols available for other researchers through its VRE portal.

The main portal is organized through several primary links: Nature - Cell Migration Gateway, migration 101, about us, cell migration knowledgebase, and CMC activity center. CMC activity center, the main section providing scientific information, knowledge, tools, and database accesses, further provides a number of links to the following topics: structure, biosensors, transgenic and knockout mice, modeling, biomaterials, imaging and photomanipulation, communications, and bioinformatics.

Table 14 Breakdown of Cluster 3 VREs based on Comprehensiveness

VRE	Total Features
Cell Migration Consortium (CMC)	23
Biomedical Informatics Research Network: Coordination Center	22
FlyBase	18
Community Climate System Model (CCSM)	17
Paleobiology Database	17
Berkeley Structural Genomics Center (BSGC)	16
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	16
Worldwide Protein Data Bank (wwPDB)	15
Visible Human Project (VHP)	14

Nature - Cell Migration Gateway links users to general information, current updates, event highlights, conference news, research library, and VRE policy. Migration 101 provides links to documents and publications. Cell migration knowledgebase provides access to databases with global and specific search options. Activity center links are mainly arranged by sub-fields. Within each sub-field, it provides further links to scientific information, protocols, publications, policies, and internal and external links to software, tools and databases. Sub-links to scientific activities/projects or output carry information about the collaborators, along with their e-mail addresses. The communication sub-link provides information and access to technology, resources, FAQ, and collaboration. The main portal page also provides a specific search option.

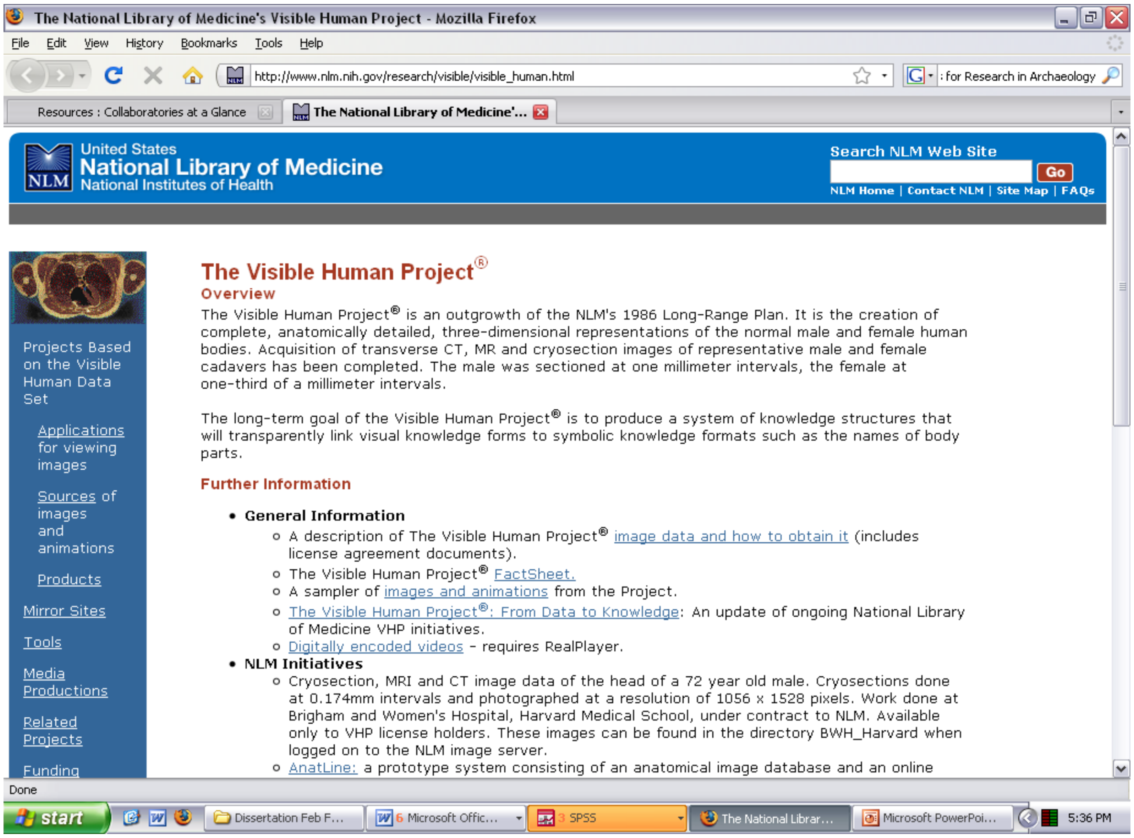
Visible Human Project

The portal of the Visible Human Project provides a primarily textual environment. This National Library of Medicine's project aims at creating a virtual environment to produce and share knowledge about human anatomy, especially, in detailed 3-D formats.

Their main portal page provides links to general information, NLM initiatives, information from the contractors for the project, proceedings from the Visible Human Project conferences, publications, send query option, information about projects based on visible human dataset, products, mirror sites, tools, media productions, related projects, and funding sources (Figure 7).

General information provides project descriptions and information and videos on project initiatives. NLM initiatives links to different sub-project activities, related information and tools, and videos. Publication leads to documents and publications resulting from the project and also to an image bank. Application and tool links provide access to software and tools needed to work within the environment.

The main page also provides a global search option and contact information for the National Library of Medicine, and a FAQ for the site. However, contact information for the VRE or search options within the VRE are totally absent.



The Visible Human Project®

Overview

The Visible Human Project® is an outgrowth of the NLM's 1986 Long-Range Plan. It is the creation of complete, anatomically detailed, three-dimensional representations of the normal male and female human bodies. Acquisition of transverse CT, MR and cryosection images of representative male and female cadavers has been completed. The male was sectioned at one millimeter intervals, the female at one-third of a millimeter intervals.

The long-term goal of the Visible Human Project® is to produce a system of knowledge structures that will transparently link visual knowledge forms to symbolic knowledge formats such as the names of body parts.

Further Information

- **General Information**
    - A description of The Visible Human Project® image data and how to obtain it (includes license agreement documents).
    - The Visible Human Project® FactSheet.
    - A sampler of images and animations from the Project.
    - The Visible Human Project®: From Data to Knowledge: An update of ongoing National Library of Medicine VHP initiatives.
    - Digitally encoded videos - requires RealPlayer.
- **NLM Initiatives**
    - Cryosection, MRI and CT image data of the head of a 72 year old male. Cryosections done at 0.174mm intervals and photographed at a resolution of 1056 x 1528 pixels. Work done at Brigham and Women's Hospital, Harvard Medical School, under contract to NLM. Available only to VHP license holders. These images can be found in the directory BWH_Harvard when logged on to the NLM image server.
    - AnatLine: a prototype system consisting of an anatomical image database and an online

Figure 7 Visible Human Project portal homepage.

A breakdown from a functionality perspective reveals five different levels within the cluster. Biomedical Informatics Research Network: Coordination Center (BIRN CC) leads the raking of functionality having all the surveyed features. Cell Migration Consortium (CMC) follows with having 12 of the 13 surveyed features. Cluster 3, in fact, contains five of the top six VREs considering functionality. The breakdown from a functionality perspective is as follows (Table 15):

Table 15 Breakdown of Cluster 3 VREs based on Functionality

VRE	Total Features
Biomedical Informatics Research Network: Coordination Center (BIRN CC)	13
Cell Migration Consortium (CMC)	12
Community Climate System Model (CCSM)	10
Paleobiology Database	10
FlyBase	10
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	10
Berkeley Structural Genomics Center (BSGC)	9
Visible Human Project (VHP)	8
Worldwide Protein Data Bank (wwPDB)	8

The breakdown from a communication perspective creates three different levels with three of the nine VREs having all the communication features surveyed, five other VREs missing just one feature, and the last one missing only two of the seven surveyed features. The ranking based on communication features is as follows (Table 16):

Table 16 Breakdown of Cluster 3 VREs based on Communication

VRE	Total Features
Cell Migration Consortium (CMC)	7
Worldwide Protein Data Bank (wwPDB)	7
Paleobiology Database	7
Biomedical Informatics Research Network: Coordination Center (BIRN CC)	6
Berkeley Structural Genomics Center (BSGC)	6
Community Climate System Model (CCSM)	6
FlyBase	6
Visible Human Project (VHP)	6
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	5

Although the presence of functionality and communication features within the VREs in Cluster 3 showed somewhat similar patterns, there was a very different picture for collaboration features. Cell Migration Consortium (CMC) came out as the forerunner with four of the five possible features closely followed by BIRN CC with three features and FlyBase with two. However, three of the VREs have only one collaborative features present and the other three have none (Table 17).

Table 17 Breakdown of Cluster 3 VREs based on Collaboration

VRE	Total Features
Cell Migration Consortium (CMC)	4
Biomedical Informatics Research Network: Coordination Center (BIRN CC)	3
FlyBase	2
Berkeley Structural Genomics Center (BSGC)	1
Community Climate System Model (CCSM)	1
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	1
Worldwide Protein Data Bank (wwPDB)	0
Paleobiology Database	0
Visible Human Project (VHP)	0

Cluster 4: Low-Function Non-Collaborative

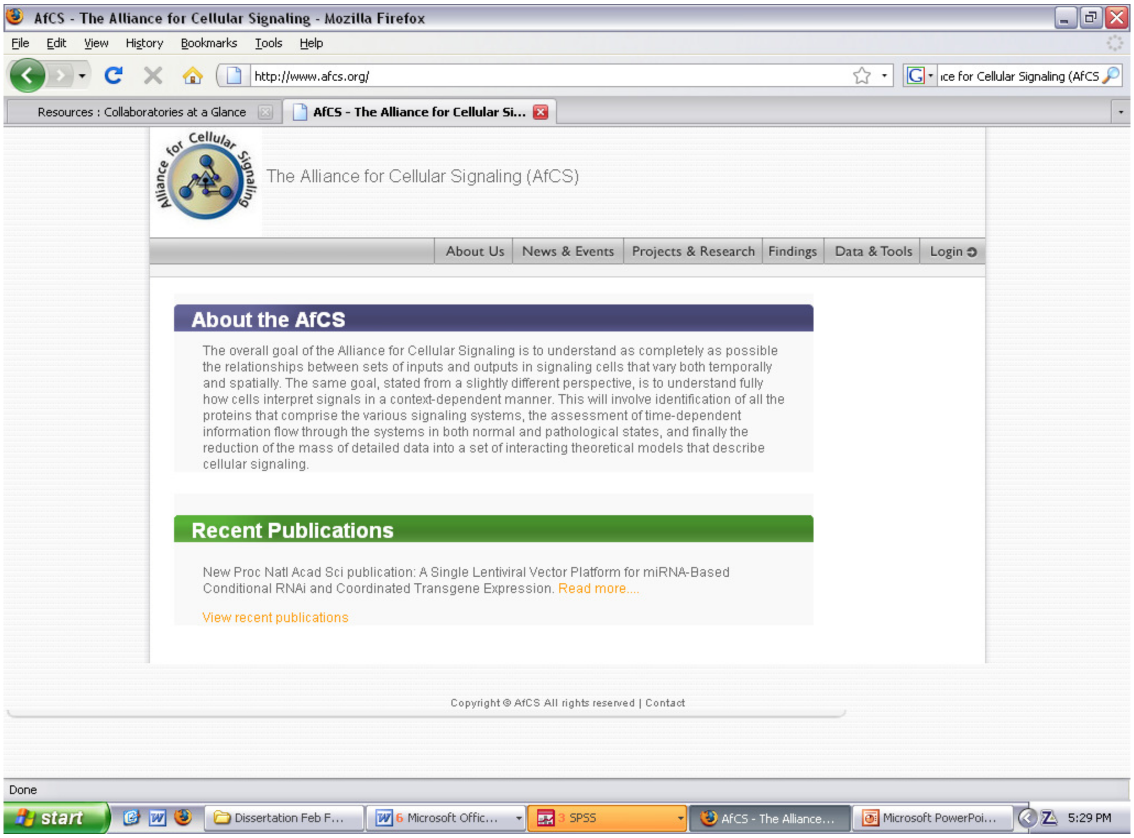
Cluster 4 constitutes the weakest set of functional VREs with low functionality, moderately low communication, and no collaboration features. Alliance for Cellular Signaling (AfCS) came out as the most comprehensive VRE within this cluster with 12 features of functionality and communication. The breakdown of the VREs from the most to the least comprehensive (based on the number of features present) are as follows (Table 18):

Table 18 Breakdown of Cluster 4 VREs based on Comprehensiveness

VRE	Total Features
Alliance for Cellular Signaling (AfCS)	12
Bugscope	10
Worm Community System (WCS)	10
Clickworkers	7

Alliance for Cellular Signaling

Though there is a moderate presence of functionality and communication features, this cluster does not contain any collaborative features. The Alliance for Cellular Signaling (AfCS) portal provides a simple and almost graphic-free web environment (Figure 8). The main page is tabbed with about the VRE, news and events, project and research, findings, data and tools, login for restricted use, and recent publications links. The main body provides a brief description of the VRE which is also the content of the about us link, however, with slightly different arrangements. About us page provides some more information about people related to the alliance without any e-mail or other contact address. It also lists alliance external laboratory URLs.



About the AfCS

The overall goal of the Alliance for Cellular Signaling is to understand as completely as possible the relationships between sets of inputs and outputs in signaling cells that vary both temporally and spatially. The same goal, stated from a slightly different perspective, is to understand fully how cells interpret signals in a context-dependent manner. This will involve identification of all the proteins that comprise the various signaling systems, the assessment of time-dependent information flow through the systems in both normal and pathological states, and finally the reduction of the mass of detailed data into a set of interacting theoretical models that describe cellular signaling.

Recent Publications

New Proc Natl Acad Sci publication: A Single Lentiviral Vector Platform for miRNA-Based Conditional RNAi and Coordinated Transgene Expression. Read more....

View recent publications

Figure 8 The Alliance for Cellular Signaling portal homepage.

The news and events link takes the user to few recent articles/news with author name and very brief description with link option to expand or view the main file. The project and research link hyperlinks to short description of projects. The findings option links to research reports, journal publications, and research slideshows. The majority of the research reports provide full text. However, journal publications provide abstract and link to the corresponding journal or journal database sites. Data and tools option provides access to public data, private data, and databases. The mentioned login option in the home page is to access private data within this environment. This VRE provides only one contact option – a global e-mail option in the homepage.

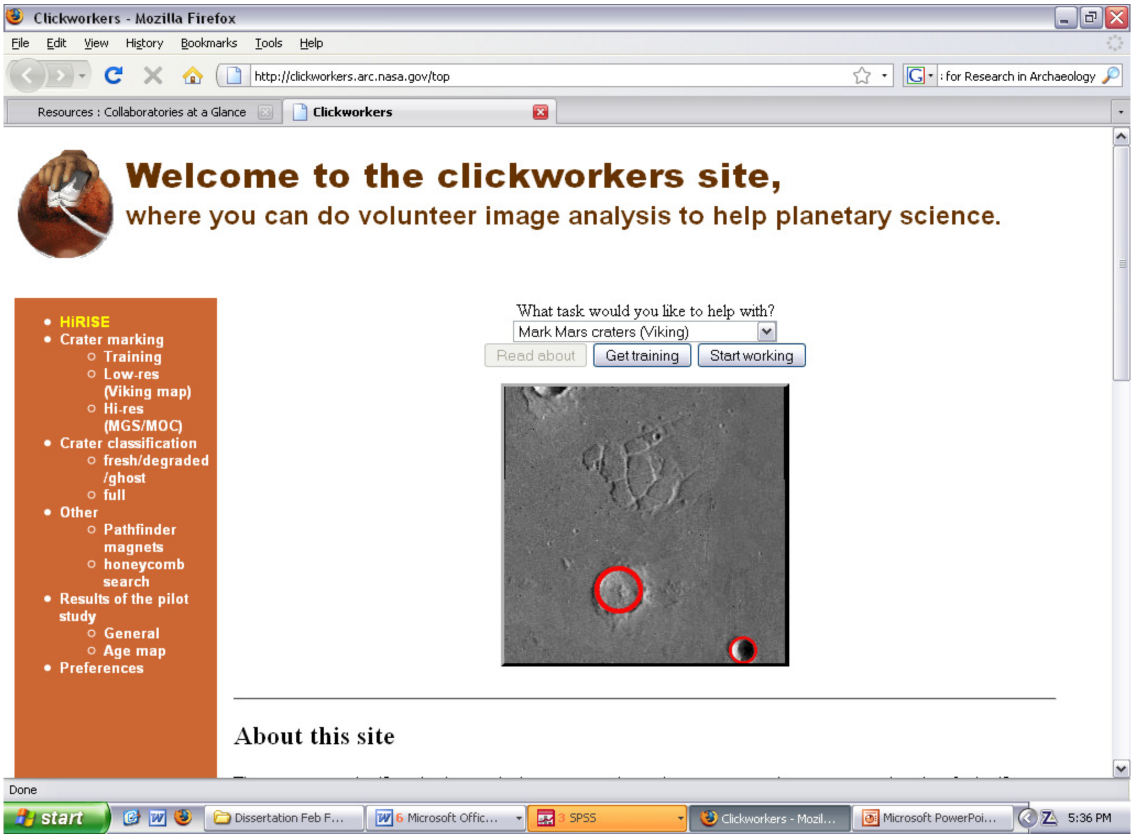
Clickworkers

Clickworkers is an open community contribution system that allows volunteers to identify craters on Mars and classify craters by age. It provides a very simple VRE portal providing link options to training, classification scheme, task choice, and start working (Figure 9). The main body also provides a description of the site and information about the original NASA project that created this VRE site. The main-body description is hyperlinked to the old main page, some research results, and more training and information option to develop ideas about crater classification related work. The homepage also contains FAQ, a privacy statement link, and a global e-mail option.

A breakdown from a functionality perspective reveals 3 different levels within the cluster (Table 19). AfCS leads the raking of functionality having 7 out of 13 surveyed features. Clickworkers bottomed the list by having only two features.

Table 19 Breakdown of Cluster 4 VREs based on Functionality

VRE	Total Features
Alliance for Cellular Signaling (AfCS)	7
Worm Community System (WCS)	6
Bugscope	5
Clickworkers	2



Welcome to the clickworkers site,

where you can do volunteer image analysis to help planetary science.

What task would you like to help with?

About this site

Figure 9 Clickworks portal homepage.

There is almost no variation in communication features within this cluster. There are only two different levels with minimum numerical difference. The ranking base on communication features is as follows (Table 20):

Table 20 Breakdown of Cluster 4 VREs based on Communication

VRE	Total Features
Bugscope	5
Alliance for Cellular Signaling (AfCS)	5
Clickworkers	5
Worm Community System (WCS)	4

This cluster mostly provides information about site, project, research, and collaborators. Clickworkers is the only VRE without collaborator information as the nature of collaborative work is voluntary. However, the same site does not provide almost any of the surveyed functions. There is a total absence of any search option throughout the cluster. Interestingly, WCS provides no e-mail link. There is also no software and only one VRE provides an analysis tool.

Cluster 5: Non-Functional Informative

The communication features of cluster 5 are better than the previous cluster. However, this cluster is almost non functional. The presence of functionality features is lowest among all clusters. Environmental Molecular Sciences Laboratory Collaboratory (EMSL) and Argonne Collaborative Access Teams scored the highest within cluster 5 by having 11 of the 25 possible features combining functionality, and communication (Table 21). Sequoia 2000, a VRE in this cluster, scored the minimum by having only one functionality (external links) and one communication (site info) feature. Sequoia 2000 reveals a very interesting VRE issue – mortality through suspended activity.

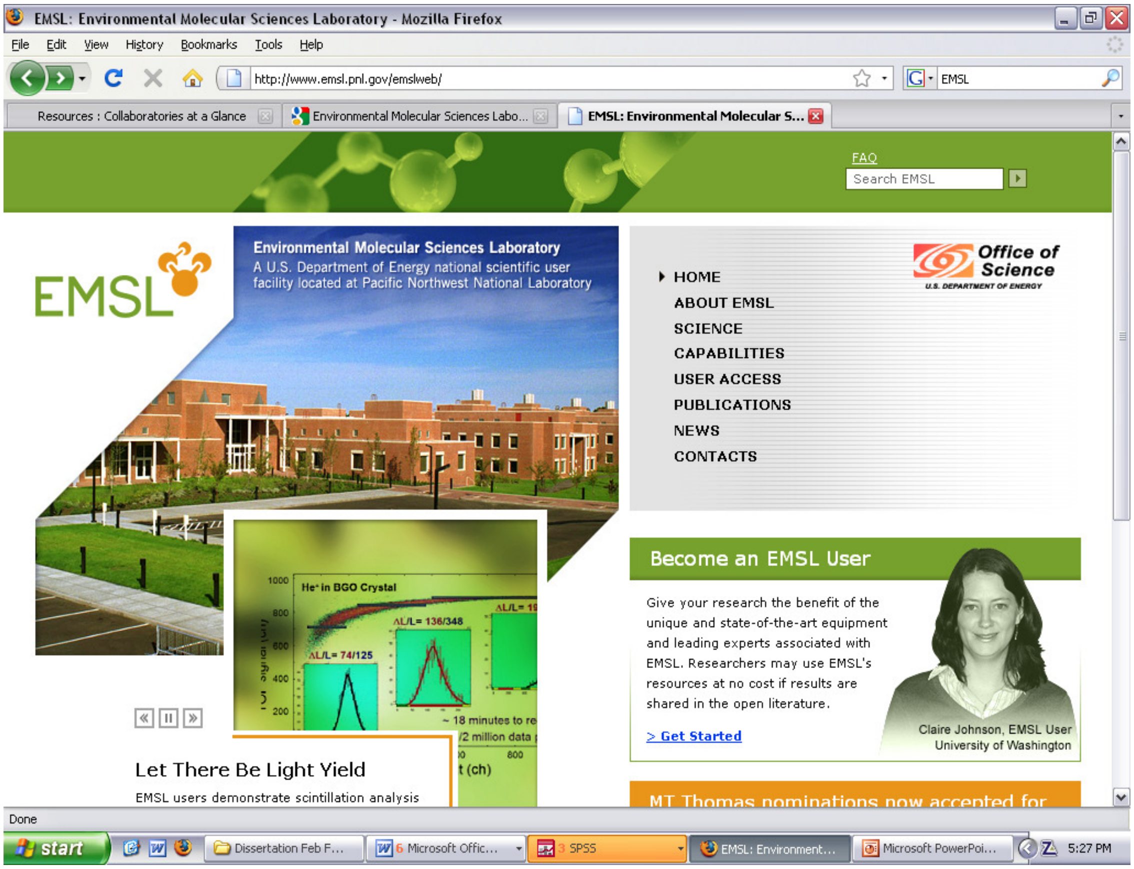
Table 21 Breakdown of Cluster 5 VREs based on Comprehensiveness

VRE	Total Features
Argonne Collaborative Access Teams	11
EMSL	11
Space Physics and Aeronomy Research Collaboratory (SPARC)	7
Materials Microcharacterization Collaboratory	6
Southern Astrophysical Research (SOAR) Telescope	3
Sequoia 2000	2

EMSL

Funded by DOE's Office of Biological Research, EMSL is a US national scientific user facility at Pacific Northwest National Laboratory. It provides integrated experimental and computational resources dedicated to environmental molecular sciences discovery. Its resources include Supercomputer, Mass spectrometers, Nuclear Magnetic-Resonance spectrometers (NMR), Surface characterization and deposition instruments, and other high-precision analytical instruments.

The main portal page provide links to About EMSL, science, capabilities, user access, publications, news, contacts, research highlights, user account information, and external links to DOE, The Biological and Environmental Research (BER), and Climate and Environmental Sciences (Figure 10). About EMSL provides brief descriptions of the collaboratory and collaborators. Science provides information projects, research, and patents. Resources link provides a full list of available instruments. User access link provides user access for restricted site, proposal submission, remote use, visit facility, and staff member related information. Publication provides access to scientific reports and brochures with specific search option and listing of documents by year. Contact allows users to find persons related to VRE along with their e-mail and telephone number. There are also FAQ and general search options in the homepage, a news section that provides VRE-related and research-interest-related news and information, and portal update information.



Environmental Molecular Sciences Laboratory
A U.S. Department of Energy national scientific user facility located at Pacific Northwest National Laboratory

Become an EMSL User

Give your research the benefit of the unique and state-of-the-art equipment and leading experts associated with EMSL. Researchers may use EMSL's resources at no cost if results are shared in the open literature.

> Get Started

Claire Johnson, EMSL User
University of Washington

Let There Be Light Yield

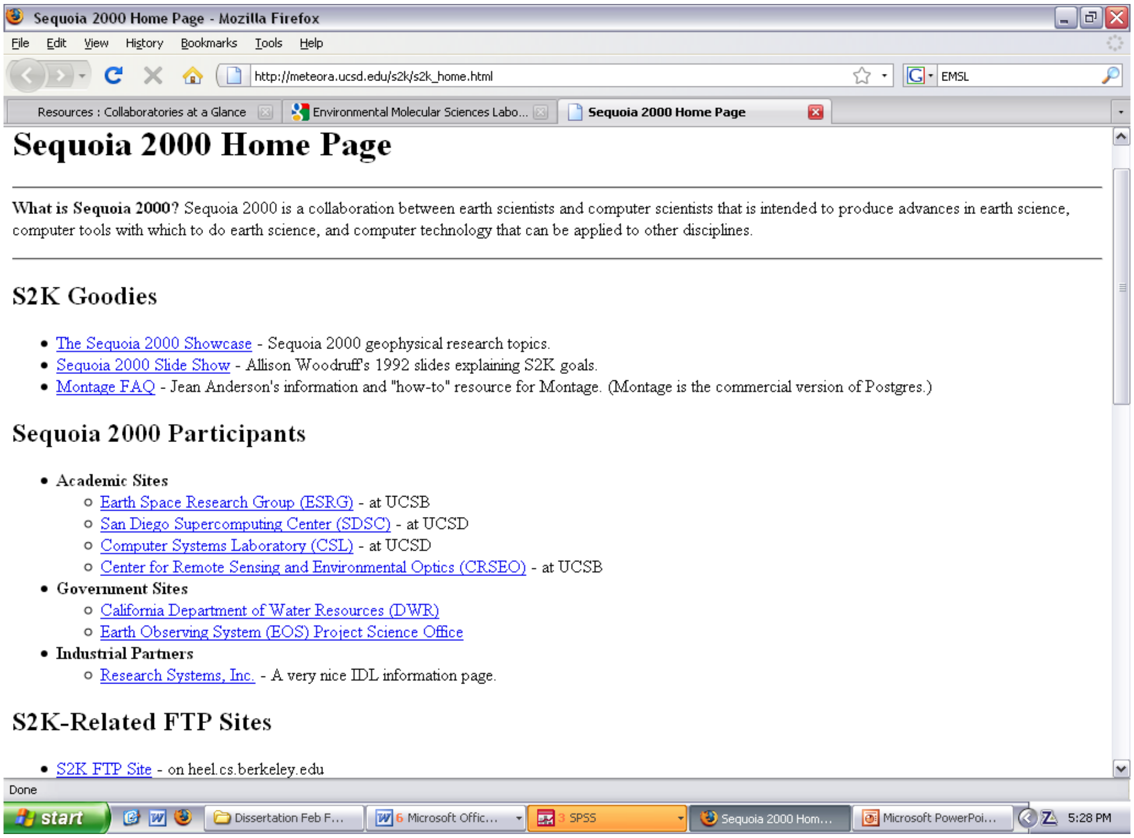
EMSL users demonstrate scintillation analysis

MT Thomas nominations now accepted for

Figure 10 EMSL portal homepage.

Sequoia 2000

Sequoia 2000 is a VRE that allows collaboration between earth scientists and computer scientists. The objective of the collaboration is to test the application of computer technology for the advancement of earth sciences, thus, also testing the fit of computer technology in multi-disciplinary scientific environments.



Sequoia 2000 Home Page

What is Sequoia 2000? Sequoia 2000 is a collaboration between earth scientists and computer scientists that is intended to produce advances in earth science, computer tools with which to do earth science, and computer technology that can be applied to other disciplines.

S2K Goodies

- The Sequoia 2000 Showcase - Sequoia 2000 geophysical research topics.
- Sequoia 2000 Slide Show - Allison Woodruff's 1992 slides explaining S2K goals.
- Montage FAQ - Jean Anderson's information and "how-to" resource for Montage. (Montage is the commercial version of Postgres.)

Sequoia 2000 Participants

- **Academic Sites**
  - Earth Space Research Group (ESRG) - at UCSB
  - San Diego Supercomputing Center (SDSC) - at UCSD
  - Computer Systems Laboratory (CSL) - at UCSD
  - Center for Remote Sensing and Environmental Optics (CRSEO) - at UCSB
- **Government Sites**
  - California Department of Water Resources (DWR)
  - Earth Observing System (EOS) Project Science Office
- **Industrial Partners**
  - Research Systems, Inc. - A very nice IDL information page.

S2K-Related FTP Sites

- S2K FTP Site - on heel.cs.berkeley.edu

Figure 11 Sequoia 2000 portal homepage.

The text-only site provides a list of links to partner institutions, research topics, and news and information (Figure 11). There are various other dead links in the site.

Altogether, any user can get brief idea about the VRE and browse some external related sites. Other than that, this site provides no research oriented, communicative, or collaborative function.

A breakdown from a functionality perspective reveals 3 different levels within the cluster. The breakdown from a functionality perspective is as follows (Table 22):

Table 22 Breakdown of Cluster 5 VREs based on Functionality

VRE	Total Features
Argonne Collaborative Access Teams	5
EMSL	5
Materials Microcharacterization Collaboratory	3
Space Physics and Aeronomy Research Collaboratory (SPARC)	3
Southern Astrophysical Research (SOAR) Telescope	0
Sequoia 2000	0

The breakdown from a communication perspective shows four different levels. The ranking based on communication features is as follows (Table 23):

Table 23 Breakdown of Cluster 5 VREs based on Communication

VRE	Total Features
Argonne Collaborative Access Teams	6
EMSL	6
Space Physics and Aeronomy Research Collaboratory (SPARC)	4
Materials Microcharacterization Collaboratory	3
Southern Astrophysical Research (SOAR) Telescope	3
Sequoia 2000	1

EMSL is the only VRE within this cluster with specific search options and analysis tools. The VREs in this cluster contains no database, downloadable software, external search options, link to external databases, tools, or VREs. EMSL and Argonne Collaborative Access Teams provide project, research and collaborator information. SPARC also provides collaborator information. There are no collaborative features present within these virtual environments. Interestingly, within this cluster Southern Astrophysical Research (SOAR) Telescope and Sequoia 2000 are purely informational sites without any functionality.

Comparison of Clusters

Results of cluster analysis show us that some commonalities forced VREs to form clusters. However, we can identify a couple of comprehensive clusters within all VREs. Though, clusters are mainly based on overall similarities among samples, comprehensiveness does not necessarily depend upon cluster. Cluster 1, 2, and 3 showed significantly stronger presence in all three aspects compared to cluster 4 and 5. Cluster 3 dominates the overall higher rankings. VREs of cluster 4 and 5 are clotted at the bottom of the list.

Comparison among cluster based on functionality shows almost the same result. Cluster 3 dominates the top ranking along with cluster 2 VREs, followed by Cluster 1 VREs are in the middle. Cluster 4 and 5 are again at the bottom of the list. Cluster 1 has a better presence in the ranking from a communication perspective. Here we observe a strong presence of cluster 1 and 2 along with cluster 3. Cluster 1 shows even a stronger presence in collaboration. However, several cluster 3 VREs are also present among the top. Cluster 2 falls behind cluster 1 and 3 in collaborative aspects.

Productivity Analysis

The analysis shows that only few VREs provide a full participant list. Most of the VREs provide a list of people related to VRE management and coordination. Often, they provide names of team leaders or principal investigator. However, information about all internal participants (note: defined in the method section) is sparse.

According to the information provided in their portal, EMSL is one of the most productive VRE. EMSL 181 scientists are working with EMSL. A total of 2173 publications came out of the VRE. The numbers of papers are even higher in Argonne lab. However, there is no specific list of participants in the Argonne Collaborative Access Team portal. However, there are project and research descriptions, sometimes, with collaborators name. The portal publication link listed a total of 7392 published articles that used the VRE facility for research. 6812 of those articles were by VRE users.

Paleobiology Database has highly organized participant information. It has 111 participants from 77 institutions. A total of 91 publications are listed in the site. Berkeley Structural Genomics Center (BSGC) provides information about 21 collaborators from 18 institutions. It lists a total of 113 publications. Biological Collaborative Research Environment (BioCoRE) portal mentions about eight-member team related to the VRE. It further talks about the theoretical and computational biophysics group members as users of the facility. A total of 39 members were listed under that group. BioCoRE mentioned about 1 conference paper and 14 other publications as a direct output of their activity. There are 26 institutions and 35 research labs affiliated with Biomedical Informatics Research Network: Coordination Center (BIRN CC). It mentions about 18 core members related mainly to coordinating activity. BIRN lists a total of 146 research publications.

Molecular Interactive Collaborative Environment (MICE) lists 2 publication and 6 participants. However, three major grids have alphabetically arranged information about documents related to that. These documents range from technical instructions to articles. 136 documents are listed in the site. Materials Microcharacterization Collaboratory (MMC) has 8 members and 5 institutional affiliations. A total of 2 publications are listed in the site. CHRONOS lists 149 participants from over 50 participating institutions. A total of 14 publications are mentioned in their portal site. The Virtual Environments for Research in Archaeology (VERA) has 18 core members from 3 participating institutions. It lists 6 papers, 2 articles, and 1 report. Community Climate System Model (CCSM) has 36 participants. CCSM is the only VRE that indicates author affiliation in their list of publications. It has a total of 350 publications. Southern Astrophysical Research (SOAR) Telescope lists a total of 4 members – a director, 2 resident astronomers, and an astronomer. It lists a total of 6 publications. Cell Migration Consortium (CMC) lists 38 participants from 20 participating institutions. A total of 301 publications are mentioned in the portal site. Collaboratory for the Multi-Scale Chemical Sciences (CMCS) lists 28 members from 9 institutions and 3 publications. Alliance for Cellular Signaling (AfCS) has 26 publications. It also provides information on participants. LIPID Metabolites and Pathways Strategy (LIPID MAPS) lists 55 members and 95 publications. Sloan Digital Sky Survey (SDSS) has 150 participants from 25 participating institutions. A total of 526 publications are listed in the site. Theoretical and Computational Biophysics Group (TCBG) lists a total of 39 participants and 494 publications.

BioModels Database does not provide information about collaborators. As the portal primarily provides models, their publication information is arranged in relation to the published models. The database lists 208 curated and 85 non-curated models. Similarly, the aim of the Virtual Cell Portal is to publish models. So far, it has published 38 models.

International Virtual Observatory Alliance (IVOA) lists 16 member organization and names of 16 principal investigators – one from each organization. It provides a list of documents like executive reports, minutes, or technology standards. However, it does not specifically mentions about research publications. BeSTGRID provides information about 9 participating institutions, a 14 member steering committee, and 16 project information with names and addresses of lead members. It also provides a list of 130 registered users. The VRE does not provide a list of scientific publication or journal articles.

FlyBase has no participant information; however, it lists 33 publications. However, FlyBase do not provide a complete author list. It uses et al in their publication that creates a major problem in authorship analysis. National Fusion Collaboratory lists 52 publications using the same method. It also lists 32 members. It also has a collaborator site that is not active.

Sequoia 2000 lists 7 institutional participants. The site does not provide any other information. Worm Community System (WCS) mentions about 7 publications but no participant information.

Visible Human Project (VHP), Worldwide Protein Data Bank (wwPDB), Space Physics and Aeronomy Research Collaboratory (SPARC), Clickworkers, and Bugscope do not provide any collaborator or publication information.

25 papers published since 2005 were sampled from each VRE that had more than 25. If the VRE had less than 25, then all papers were included in the sample. VREs that did not have publications after 2004 were not included in the sample. Altogether, 15 VREs provided enough information to determine degree of internal-external collaboration (Table 24).

Table 24 Sampled VREs with Author and Collaboration Information

VRE Name	Sample Size	Total Author	Internal Authors	Internal-External Collaborative Papers
EMSL	25	75	16	13
Virtual Cell Portal	23	75	7	8
Argonne Collaborative Access Teams	25	59	21	13
CHRONOS	7	19	5	2
VERA	9	9	9	0
SOAR Telescope	5	26	0	0
CMC	25	71	16	24
BIRN CC	9	34	7	3
BSGC	25	34	7	23
AfCS	11	66	24	8
CCSM	25	69	25	11
CMCS	1	26	26	0
Paleobiology Database	25	35	22	11
LIPID MAPS	25	47	16	25
TCBG	25	40	13	19

The majority of the VREs with productivity information (almost 70%) came from cluster 1 and cluster 3 (Table 25). Both these clusters scored higher than other clusters in comprehensiveness. EMSL and Argonne Lab somewhat acts as outliers. Both the VRE went through some changes in between primary and secondary data collection phases. The analysis excluded these VREs from the list.

Table 25 Authorship and Collaboration Information by Cluster

Cluster	Frequency	Percent
Full-service Balanced	4	26.7
Simple Non-collaborative	1	6.7
Full-service Functional	6	40.0
Low-function Non-collaborative	1	6.7
Non-functional Informative	3	20.0
Total	15	100.0

The analysis shows us that, there is a significant correlation between the total number of publication and comprehensiveness, $r = .48$, p (one tailed) $< .01$. Among features, there is a significant correlation between the total number of papers and functionality, $r = .49$, p (one tailed) $< .01$, and communication, $r = .34$, p (one tailed) $<$ .05. Among functionality features, total number of publication has significant correlation with databases ($r = .4$), global search options ($r = .42$), specific search options ($r = .38$), and external search options (*r* = .53), *p* (one tailed) < .05. Total number of publication also has significant correlation with collaborators information (*r* = .32) and specific e-mail (*r* = .36), *p* (one tailed) < .05.

There is a significant relationship between the percentage of collaborative papers and comprehensiveness, *r* = .38, *p* (one tailed) < .05. Among the features, functionality is correlated with the percentage of collaboration paper, *r* = .44, *p* (one tailed) < .01. Among functionality features, specific search options (r = .36), link to external databases (*r* = .32), link to external tools (*r* = .52), and link to external VREs (*r* = .47) is significantly correlated with percentage of collaborative publications, *p* (one tailed) < .05. Among other features, newsgroup option was correlated with percentage of collaborative publications, *r* = .4, *p* (one tailed) < .05. Interestingly, presence of a FAQ was negatively correlated with percentage of collaborative publications, *r* = -.32, *p* (one tailed) < .05.

Though percentage of external authors was not correlated with any features, it was correlated with specific search options (*r* = .34), search option for external sources (*r* = .36), and link to external tools (*r* = .32), *p* (one tailed) < .05. Again, FAQ was negatively correlated with the percentage of external authors, *r* = -.43, *p* (one tailed) < .05.

CHAPTER V

DISCUSSION AND CONCLUSIONS

This study investigated the technological arrangements of virtual research environments (VRE) and the relationship between technological features and effectiveness. VRE portals were at the core of the investigation as they are the entry points for VRE related information and resource access. VRE portals not only provide links to resources within that environment, but also influence the perception of a user about VRE effectiveness. Effectiveness in this study was measured as productivity. Productivity was viewed as a combination of the amount of publications produced by a particular VRE (as reported by them), presence of the external authors in the publication, and co-authorship of papers by both internal members of the VRE and external authors who are not necessarily a part of the VRE.

Technological Configurations: Factor behind a New Classification

This research shows that the technological arrangements of the VRE neither depend upon scientific discipline nor the existing functional typology. A typology developed based on the technological arrangements in VREs identified five different types of VRE systems. These types were different from those in the Bos et al.'s (2007) typology.

The study did not identify a significant presence of communication and collaboration technologies within the VRE systems. The majority of the VREs did not have any systematic arrangement of collaboration technologies. The presence of communication features were in fact much higher than collaboration features. Case studies show that research, project, and collaborator information were not properly structured to meet users' needs. However, there were exceptions. Cluster analysis revealed that cases were clustered when there was a significant presence of communication or collaboration technologies.

There is a balanced mixture of technological features within some VRE systems. These VREs, labeled as "Full-service Balanced" and "Full Service Functional," provide features from all three subcategories of attributes – functionality, communication, and collaboration. However, "Full Service Functional" VREs provide extensive features supporting scientific research. The majority of the VREs (17 out of 31) fall under these categories. There are four more "Simple Non-collaborative" VREs providing moderate functionality but no collaborative aspects. The result shows us that functionality is a major focus in the virtual environments. Two- thirds of the VREs include high to moderate functional aspects.

Technological arrangements within VREs, therefore, played a major role in the classification scheme. It justified the importance to technology. The previous classification did not neglect the role of technology. However, they downplayed the significance of technological features.

Bos et al. (2007) mentioned that their project had conducted a technology inventory of the collaboratories. However, developing a taxonomy based on technology was not their objective. Their classification system

> sought to identify organizational patterns, somewhat similar to design patterns .. which could be used by funders and project managers in designing new collaborations. Rather than focusing on the technology or the emergent organizational features, the scheme is tightly focused on the goals of the projects. The result of this classification should be identification of key challenges and recommendation of practices, technology, and organizational structures that are appropriate for a stated set of goals. P. 656

Their research project also asked the question ‘‘What technology should be recommended for collaboratories?” (p.669). However, they believed that the question could not be addressed because of the diverse nature of practices related to VREs. The authors argued that the technology needs of one type of VRE should be fundamentally different from other types. However, this research has a different focus and the findings are not exactly in line with their predictions.

Our cluster analysis suggests that the technology implementation within VREs followed specific patterns. However, technology configurations of VREs are not fundamentally similar within Bos et al.’s category or fundamentally different across categories. There are similarities and differences within and across their categories. For this reason, technological feature based survey placed VREs of different types of Bos et al.’s classification into the same cluster.

This situation raises two different questions. First, how are VREs implementing technologies? Second, it forces us to rethink the question of recommending a proper technological configuration for VREs. The study shows that the comprehensiveness of a system has a relationship with its success. The technological configurations of relatively comprehensive system were dominated primarily by functionality features and sometimes by a mixture of functionality and communication features. The results of this study reveal several dominant functionality and communication features. Databases, search options, links, e-mail, collaborator information, and newsgroup options came out as very important features. Most of the successful VREs, in the absence of collaborative technology, systematically arranged the mentioned features to become productive.

Productivity: Its Relationship with Functionality, Communication, and Collaboration

Our analysis shows that seven of the VREs did not have any publications. Three other VREs had up to three publications. Therefore, 21 VREs had some publication listed as output of VRE related research. However, when we sampled papers from 2005 to 2009, we lost another six VREs as they either did not have any papers published after 2004 or their publications were not scholarly. We can also predict that several VREs did not list their papers. BeSTGRID, for example, was one of the neatly arranged VRE without any scholarly publications listed. As the aim of the research was to explore the portal, any information not provided on the portal influenced the way we look at productivity. It is because of this that the aim of the research was to look at VREs from the user’s perspective. Portals are the primary entry points for any user. Lack of publication information in the portal would force a user to think negatively about the productive nature of that environment. It will also affect the functionality, as it hampers knowledge transfer through literature depository.

Among the fifteen VREs with publication from 2005, only nine VREs had a significant or moderate number of publications. Therefore, throughout the sample we observe a low amount of publications coming out of those VREs.

Comprehensiveness of a VRE was correlated with both the number of publications by the members of a VRE and the number of collaborative papers coming out of a VRE. However, the effect of comprehensiveness was mainly dependent upon functionality. Communication and collaboration aspects, as predicted in the communication and information literatures did not show the impact on productivity as they were supposed to.

Productivity & Functionality

Search options came out as the most significant features in VRE effectiveness. Global, specific, and external search options were significantly correlated with the total number of publications. Specific search option influenced collaborative paper output and the amount of external authors. The importance of search options points us to two different aspects.

Data analysis shows that, those who actually participated in the environment needed to use search options to find information. We can argue that the arrangement of information and resources within the VRE portals were not easy enough for them to explore, and therefore, they needed to use various search options. It is also true for this research. As the participant list within the VRE was not readily available in most cases, we needed to use the search option within the system to discover if the name appeared in the list of author was of a member of the VRE. Any VRE user might go through the name problem to identify concepts, research, or collaborator information.

Beside search options, external links were the other features of importance. External links were especially for collaborative paper and external authors. Links to external databases, tools, and VRE significantly influenced internal-external co-authorship. Links to external tools was also related to the number of external authors. How might we explain this?

External links provide a linkage between two VREs, thereby creating a network between two groups of users. This link helped them to know about each other, the line of research or projects the other VRE was conducting, and the research expertise available outside their own environment. This option enabled the development of a scientific collaborative network. It might also shed light on the importance of the search options. Specific search options were significantly correlated with external databases and tools. External search options were related to databases and external VREs. Therefore, the presence of external links influenced users to explore more information related to a topic or a member of the VRE as that information was not readily available. It was also noted that the search option was essential in data mining. Having a database itself was significantly correlated with the total number of publication.

Productivity & Communication

Communication as sub category was only significantly correlated with the total number of publications. Collaborator information and specific e-mail addresses were important features influencing publication index. However, research information played a role in collaborative papers. It was obvious that users were looking for information about research, participating researchers, and a way to communicate with them. The absence of communication features, therefore, explains the low number of publications observed in the sample.

Very interestingly, FAQ as a communicative feature has a negative correlation with both the number of external authors and the number of collaborative publication. Maybe FAQ is useful in normal web sites but not perceived as a useful feature in scientific portals.

The negative importance of FAQ may be explained by the hypothesis that a scientific audience does not like to go through a list of questions and answers provided to them. Rather, they like to explore the topic of their interest and also to solve problem related to VRE. It might support the idea that the scientists are comparatively more individualistic than others as higher education places more emphasis on individual accomplishments (Mohrman et al., 1995; Shannon, 1980) and they possess a self-directive work ethic (Hackman, 1990). However, the search option reflects their desire for challenging work and interest in knowledge sharing – the factor that makes them collaborative (Hackman, 1990). FAQ was also related mostly to project information and conferencing systems. FAQ either provided project related information or "how to" instruction for conferencing.

Productivity & Collaboration

Newsgroup was the only collaboration feature significantly related to productivity. Thirteen of the surveyed VREs had newsgroups. The presence of newsgroup was positively correlated with collaborative publication. Other collaborative features were not significant as all of those had a very low presence in the sample. We only had three teleconferencing, three wikis, five chat-rooms, and five blogs within the sampled VREs.

Concluding Remarks: Some Thoughts about Technology and Future VRE Developments

This research highlights the lack of collaborative technologies within VREs. Though the research does not link collaborative technologies to the effectiveness of the VRE based on our productivity measure, it poses a couple of significant questions. Let us summarize the observed scenario before exploring those questions.

First, we observe a tendency of network development within scientists using the available technology. Scientific productivity was higher when they could employ search options, when the VRE actually exposed them to external collaborators and resources, or when there was collaborator information present within the system. There was also a tendency to appropriate these features for scientific productivity. Second, newsgroup, the only collaborative feature with somewhat significant presence, came out as an influential factor. Third, specific-email influenced the number of scientific publications. Finally, we should remember that the overall productivity level was low.

So the first question is: what would be the impact of collaboration tools? The research shows us that collaboration tools affect productivity. However, these tools are not available within the systems studied in this research. The literature on e-collaboration identified the significance of different collaboration tools. For example, research shows that, when present, a weblog can be very influential for learning (Luzon, 2008). Finding like minded people, asking for advice or receiving feedback, and developing virtual community based on discussions are three most influential outcome of academic weblog (Luzon, 2008). In the research, we have seen users performing these functions through other features solely because of availability.

The second question is: why were communication features not significantly related to VRE effectiveness? Most of the communication features were present within the sampled VREs. The only feature that had a weaker presence was specific-email system. Again, it forces us to think about the criteria we are using to implement technology within VRE. Although, most of the communication features were present, the case studies tell us that the arrangement of those technologies were often not proper. Moreover, presence of a communication feature does not necessarily guarantee usability. Research and collaborator information were perfect example of that. Most of the VREs had some information about participants within that system and the type of research they are engaged with. However, very seldom did they provide sufficient information that allows a user to develop a proper idea about research activities, the people related to those activities along with their task division and expertise, and a system to communicate to those people. However, whenever such information was available, it was significantly related to productivity.

Our literature review argues that there is no specific guideline in implementing technology. Our analysis reveals the important aspects of several technological features. We can also argue that there is logic behind the importance of those technologies. Based on our analysis, we can draw the following conclusions:

- Functionality is an essential part of any VRE. Functionality is correlated with VRE effectiveness.
- Communication and collaboration technologies influence the effectiveness of VRE systems. The absence of such technology leads to lower productivity.
- VREs need to place more emphasis in adding collaboration technologies within their systems.
- Information within a VRE system needs to be structured, clear, and complete. Project and research information should accompany collaborator information along with an option enabling a direct communication to those collaborators.
- Any VRE should include a comprehensive search system. Our result shows us that VRE participants have a comparatively higher tendency to use search systems.
- VREs need to link them within similar VREs, projects, or other sites. There is a tendency among the VRE users to use these links to explore potential collaborators.

This research supports the idea that communication and collaboration technologies are important in VRE effectiveness. Absence of those technologies contributes to low productivity. However, it also tells us that such technologies are not yet available or properly implemented within those systems. Future VRE development should pay more attention in technological arrangements. VREs also need to rethink their design process. A neat arrangement is always useful for a new user to explore VRE information. We have observed some graphics-only sites during our case studies.

Developer should note that graphics can add an aesthetic quality to a website. However, these are minor considerations.

Limitations of the Study

This study was conducted over a specific period of time. Within this time period several sites went through some changes. That is a reality for any portal based investigation. A few websites also went offline during these time period. They still have an online presence, but they are not functioning. Worm Community System and Sequoia 2000 are examples. This study did not exclude them from the list because mortality has been observed as a related attribute for VREs. Several VREs like EMSL improved during this period. The numbers of publication they have in fact reflect the improvement. Therefore, any survey using the same method and with the same sample may not provide the same result.

One of the major problems in obtaining data was related to collaborator information. Most of the sites did not provide collaborator information with proper organization. Several sites did not provide a communication link to the collaborators. In most cases, management and coordination personnel information were available. In some cases project information included the name of project leaders. In only a very few cases was a total list of participants provided. The lack of information about VRE participants was an obstacle in determining internal-external collaboration.

This study relied upon the information provided in the web portals. The objective of the research was to conduct an investigation based on provided information. The study hypothised that impression of a VRE, from a user perspective, would depend upon provided information. Therefore, the study does not claim that the information was exact or complete. There is also an ongoing rapid development in the field of information and communication technologies. Any future study, therefore, should conduct a primary survey to modify the existing feature list.

REFERENCES

Awre, C. & Ingram, C. (2005). CREE feasibility study on presenting communication and collaboration tools within different contexts. Retrieved May 25, 2007, from http://www.hull.ac.uk/cree/downloads/CREEcommsresults.pdf.

Barge, J. K. (2002). Enlarging the meaning of group deliberation: From discussion to dialogue. In Frey, L. R. (Ed.), *New directions in group communication* (pp. 159-178). Thousand Oaks, CA: Sage.

Benford, S., Greenhalgh, C., Rodden, T. & Pycock, J. (2001). Collaborative virtual environments. *Communications of the ACM, 44*(7), 79-86.

Berelson, B. & Sills, D. L. (1960). *Review of studies in the flow of information among scientists*. Columbia University: Bureau of Applied Social Research.

Bos, N.D., Zimmerman, A., Olson, J., Yew, J., Yerkie, J., Dahl, E. & Olson, G. (2007). From shared databases to communities of practice: A taxonomy of collaboratories. *Journal of Computer Mediated Communications, 12*(2), 652–672.

Carley, K. & Wendt, K. (1991). Electronic mail and scientific communication: A study of the soar extended research group. *Knowledge: Creation, Diffusion, Utilization, 12*(4), 406-440.

Chin, G., Myers, J., & Hoyt, D. (2002). Social networks in the virtual science laboratory. *Communications of the ACM, 45*(8), 87-92.

DeSanctis, G., D'Onofrio, M. J., Sambamurthy, V., & Poole, M. S. (1989). Comprehensiveness and restrictiveness in group decision heuristics: Effects of computer support on consensus decision making. *Proceedings of the 10th International Conference on Information Systems.* Boston, Massachusetts.

DeSanctis, G. L., & Poole, M. S. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science, 5*(2), 121-147.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17*, 397-434.

Ellis, C. A., Gibbs, S. J., & Rein, G. (1991). Groupware: Some issues and experiences. *Communications of the ACM, 34*(1), 39-58.

Esler, S. L. & Nelson, M. L. (1998). Evolution of scientific and technical information distribution. *Journal of the American Society for Information Science, 49*(1), 82-91.

Finholt, T. A. & Olson, G. M. (1997). From laboratories to collaboratories: A new organizational form for scientific collaboration. *Psychological Science, 8*(1), 28-36.

Fraser, M. A. (2005, July). Virtual Research Environments: Overview and Activity. *Ariadne, 44*. Retrieved April 16th, 2006, from http://www.ariadne.ac.uk/issue44/fraser/.

Gan, G., Ma, C., & Wu, J. (2007). *Data clustering: Theory algorithms, and applications*. Alexandria, VA: SIAM.

Gladstein, D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative Science Quarterly, 29*(4), 499-517.

Guzzo, R. A. & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology, 47*, 307–338.

Hackman, J. R. (1990). *Groups that work (and those that don't): Creating conditions for effective teamwork.* San Francisco, CA: Jossey-Bass.

Hagstrom, W. O. (1964). Traditional and modern forms of scientific teamwork. *Administrative Science Quarterly, 9*(3), 241-263.

Hey, T. & Trefethen, A. E. (2005). Cyberinfrastructure for e-Science. *Science, 308*, 817-821.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From Input-Process-Output models to IMOI models. *Annual Review of Psychology, 56*, 517-543.

Katzenbach, J.R. & Smith, D.K. (1993). *The wisdom of teams: Creating the high-performance organization*. Boston: Harvard Business School.

Klahr, D. & Simon, H. A. (2001). What have psychologists (and others) discovered about the process of scientific discovery? *Current Directions in Psychological Science, 10*, 75-79.

Kouzes, R. T. (not dated). Collaboratories: Scientists working together apart. Retrieved April 16th, 2006, from http://collaboratory.emsl.pnl.gov/resources/publications/WorkingTogetherApart/world.html.

Kouzes, R. T., Myers, J. D., & Wulf, W. A. (1996). Collaboratories: Doing science on the Internet. *Computer, 29*(8), 40-46.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7*(3), 77-124.

Kraut, R. E., Egido, C. & Galegher, J. (1990). Patterns of contact and communication in scientific research collaboration. In Galegher, J., Kraut, R. E. & Egido, C (Eds.), *Intellectual teamwork: Social and technological foundations of cooperative work* (pp. 149-171). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lievrouw, L. A. & Carley, K. (1990). Changing patterns of communication among scientists in an era of “Telescience.” *Technology in Society, 12*, 457-477.

Luzon, M. J. (2008). Academic weblogs as tools for e-collaboration among researchers. In Kock, N. (Ed.), *Encyclopedia of E-collaboration* (pp. 1-6). Hershey, NY: Information Science Reference.

MacEachren, A. M. (2001). Cartography and GIS: Extending collaborative tools to support virtual teams. *Progress in Human Geography, 25*(3), 431–444.

Mikhailov, A. I., Chernyia, A. I., & Giliarevskii, R. S. (1984). *Scientific communication and informatics*. Arlington, VA: Information Resources Press.

Mohrman, S. A., Cohen, S. G. & Mohrman, A. M., Jr. (1995). *Designing team-based organizations*. San Francisco, CA: Jossey-Bass.

Moreland, R. L. & Levine, J. M. (1988). Group dynamics over time: Development and socialization in small groups. In McGrath, J. E. (Ed.), *The psychology of time* (pp. 151-181). Newbury Park: Sage.

Nielsen, J. & Mack, R. L. (1994). *Usability inspection methods*. New York: John Wiley & Sons.

Okada, T. & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science, 21*(2), 109-146.

Olsina, L., Godoy, D., Lafuente, G. J. & Rossi, G. (1999). Specifying quality characteristics and attributes for Websites. Paper presented at the First ICSE Workshop on Web Engineering (WebE-99), Los Angeles, USA.

Olson, G. M. (2004). Collaboration Research & Collaboration Readiness. Retrieved May17, 2007, from www.crito.uci.edu/eco-workshop/olson.pdf.

Paré, G., & Dubé, L. (1999). Virtual teams: An exploratory study of key challenges and strategies. *Proceedings of the 20th International Conference on Information Systems, ICIS, 13-15th December, 1999*, Charlotte, NC.

Pelz, D. C. & Andrews, F. M. (1976). *Scientists in organizations: Productive climates for research and development*. Ann Arbor, MI: Institute for Social Research, The University of Michigan.

Polys, N. F., Bowman, D. A., & North, C. (2004). Information-rich virtual environments: Challenges and outlook. Retrieved April 16th, 2006, from http://www.domos.be/overal/white_papers/Polys_et_al.pdf.

Poole, M. S. & Ahmed, I. (in press). Media as communication: 62. Group decision systems. In Eadie, W. F. (Ed.), *21st century communication: A reference handbook*. Thousand Oaks, CA: Sage.

Poole, M. S. & Baldwin, C. L. (1996). Developmental processes in group decision making. In Hirokawa, R. Y. & Poole, M. S. (Eds.), *Communication and group decision making* (2nd ed., pp. 215-241). Thousand Oaks, CA: Sage.

Poole, M. S. & Roth, J. (1989). Decision development in small groups IV: A typology of group decision paths. Human Communication Research, 15(3), 323-356.

Poole, M. S. & Zhang, H. (2005). Virtual teams. In S. Wheelan (Ed.) *The handbook of group research and practice* (pp. 363-384). Thousand Oaks, CA: Sage

Recardo, R. & Jolly, J. (1997). Organizational culture and teams. *SAM Advanced Management Journal, 62*(2), 4-7.

Rice, J. R. & Boisvert, R. F. (1996). From scientific software libraries to problem-solving environments. *IEEE Computational Science & Engineering, 3*(3), 44-53.

Schunn, C. D. & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science, 23*(3), 337-370.

Schunn, C. D., & Klahr, D. (1995). A 4-space model of scientific discovery. *Proceedings of the 17th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Schur, A. Keating, K. A., Payne, D. A., Valdez, T., Yates, K. R., & Myers, J. D. (not dated). Collaborative suites for experiment-oriented scientific research. Retrieved April 16th, 2006, from http://collaboratory.emsl.pnl.gov/presentations/papers/ACMpaper.html.

Shannon, R.E. (1980). *Engineering management*. New York: John Wiley & Sons.

Simon, H. A. (1973). Does scientific discovery have a logic? *Philosophy of Science, 40*(4), 471-480.

Simon, H.A., & Lea, G. (1974) Problem solving and rule induction: A unified view. In Gregg, L. W. (Ed.), *Knowledge and cognition*. Hillsdale, NJ: Erlbaum.

Springmeyer, R. R., Blattner, M. M., & Max, N. L. (1992). A characterization of scientific data analysis process. *Proceedings of the 3rd Conference on Visualization '92, October 19-23, 1992*, Boston, MA.

Thylefors, I., Persson, O., & Hellström, D. (2005). Team types, perceived efficiency and team climate in Swedish cross-professional teamwork. *Journal of Interprofessional Care, 19*(2), 102-114.

Walsh, J., P. & Bayma, T. (1996). The virtual college: Computer-mediated communication and scientific work. *The Information Society, 12*(4), 343 – 363.

Wilbur, S. (1997). Models and metaphors for video mediated communication. In Finn, K. E., Sellen, A. J., & Wilbur, S. B. (Eds.), *Video-mediated communication* (pp. 351-362). Mahwah, NJ: Lawrence Erlbaum.

Zhao, H. & Georganas, N. D. (2001). Collaborative virtual environments: Managing the shared spaces. *Networking and Information Systems Journal, 10*(10), 1-23.

Zhuge, H. (2005). The future interconnection environment. *Computer, 38*(4), 27-33.

APPENDIX A

RANKINGS OF VRES BASED ON TOTAL FEATURES

VRE Name	Cluster	Total Features
Cell Migration Consortium (CMC)	3	23
BIO Informatics Research Network: Coordination Center	3	22
BeSTGRID	1	19
BioModels Database	2	18
FlyBase	3	18
Community Climate System Model (CCSM)	3	17
Paleobiology Database	3	17
Sloan Digital Sky Survey (SDSS)	2	17
Molecular Interactive Collaborative Environment (MICE)	2	16
CHRONOS	1	16
BIO Collaborative Research Environment (BioCoRE)	1	16
Berkeley Structural Genomics Center (BSGC)	3	16
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	1	16
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	3	16
Theoretical and Computational Biophysics Group (TCBG)	1	16
Worldwide Protein Data Bank (wwPDB)	3	15
International Virtual Observatory Alliance (IVOA)	1	15
National Fusion Collaboratory	1	14
Visible Human Project (VHP)	3	14
Virtual Cell Portal	2	13
The Virtual Environments for Research in Archaeology (VERA)	1	13
Alliance for Cellular Signaling (AfCS)	4	12
EMSL	5	11
Argonne Collaborative Access Teams	5	11
Bugscope	4	10
Worm Community System (WCS)	4	10
Space Physics and Aeronomy Research Collaboratory (SPARC)	5	7
Clickworkers	4	7
Materials Microcharacterization Collaboratory	5	6
Southern Astrophysics Research (SOAR) Telescope	5	3
Sequoia 2000	5	2

APPENDIX B

RANKINGS OF VRES BASED ON FUNCTIONALITY FEATURES

VRE Name	Cluster	Number of Features
BIO Informatics Research Network: Coordination Center	3	13
Cell Migration Consortium (CMC)	3	12
BioModels Database	2	11
FlyBase	3	10
Community Climate System Model (CCSM)	3	10
Paleobiology Database	3	10
Sloan Digital Sky Survey (SDSS)	2	10
Molecular Interactive Collaborative Environment (MICE)	2	10
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	3	10
BeSTGRID	1	9
Berkeley Structural Genomics Center (BSGC)	3	9
Virtual Cell Portal	2	9
CHRONOS	1	8
BIO Collaborative Research Environment (BioCoRE)	1	8
Theoretical and Computational Biophysics Group (TCBG)	1	8
Worldwide Protein Data Bank (wwPDB)	3	8
Visible Human Project (VHP)	3	8
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	1	7
Alliance for Cellular Signaling (AfCS)	4	7
International Virtual Observatory Alliance (IVOA)	1	6
Worm Community System (WCS)	4	6
National Fusion Collaboratory	1	5
The Virtual Environments for Research in Archaeology (VERA)	1	5
EMSL	5	5
Argonne Collaborative Access Teams	5	5
Bugscope	4	5
Space Physics and Aeronomy Research Collaboratory (SPARC)	5	3
Materials Microcharacterization Collaboratory	5	3
Clickworkers	4	2
Southern Astrophysics Research (SOAR) Telescope	5	0
Sequoia 2000	5	0

APPENDIX C

RANKINGS OF VRES BASED ON COMMUNICATION FEATURES

VRE Name	Cluster	Number of Features
Cell Migration Consortium (CMC)	3	7
BioModels Database	2	7
Paleobiology Database	3	7
Sloan Digital Sky Survey (SDSS)	2	7
BeSTGRID	1	7
BIO Collaborative Research Environment (BioCoRE)	1	7
Worldwide Protein Data Bank (wwPDB)	3	7
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	1	7
International Virtual Observatory Alliance (IVOA)	1	7
National Fusion Collaboratory	1	7
BIO Informatics Research Network: Coordination Center	3	6
FlyBase	3	6
Community Climate System Model (CCSM)	3	6
Molecular Interactive Collaborative Environment (MICE)	2	6
Berkeley Structural Genomics Center (BSGC)	3	6
CHRONOS	1	6
Theoretical and Computational Biophysics Group (TCBG)	1	6
Visible Human Project (VHP)	3	6
EMSL	5	6
Argonne Collaborative Access Teams	5	6
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	3	5
Alliance for Cellular Signaling (AfCS)	4	5
Bugscope	4	5
Clickworkers	4	5
Virtual Cell Portal	2	4
Worm Community System (WCS)	4	4
The Virtual Environments for Research in Archaeology (VERA)	1	4
Space Physics and Aeronomy Research Collaboratory (SPARC)	5	4
Materials Microcharacterization Collaboratory	5	3
Southern Astrophysics Research (SOAR) Telescope	5	3
Sequoia 2000	5	1

APPENDIX D

RANKINGS OF VRES BASED ON COLLABORATION FEATURES

VRE Name	Cluster	Number of Features
Cell Migration Consortium (CMC)	3	4
BeSTGRID	1	3
BIO Informatics Research Network: Coordination Center	3	3
The Virtual Environments for Research in Archaeology (VERA)	1	3
Collaboratory for the Multi-Scale Chemical Sciences (CMCS)	1	2
International Virtual Observatory Alliance (IVOA)	1	2
National Fusion Collaboratory	1	2
FlyBase	3	2
CHRONOS	1	2
Theoretical and Computational Biophysics Group (TCBG)	1	2
BIO Collaborative Research Environment (BioCoRE)	1	1
Community Climate System Model (CCSM)	3	1
Berkeley Structural Genomics Center (BSGC)	3	1
LIPID Metabolites And Pathways Strategy (LIPID MAPS)	3	1
BioModels Database	2	0
Paleobiology Database	3	0
Sloan Digital Sky Survey (SDSS)	2	0
Worldwide Protein Data Bank (wwPDB)	3	0
Molecular Interactive Collaborative Environment (MICE)	2	0
Visible Human Project (VHP)	3	0
EMSL	5	0
Argonne Collaborative Access Teams	5	0
Alliance for Cellular Signaling (AfCS)	4	0
Bugscope	4	0
Clickworkers	4	0
Virtual Cell Portal	2	0
Worm Community System (WCS)	4	0
Space Physics and Aeronomy Research Collaboratory (SPARC)	5	0
Materials Microcharacterization Collaboratory	5	0
Southern Astrophysics Research (SOAR) Telescope	5	0
Sequoia 2000	5	0

VITA

Name: Iftekhar Ahmed

Address: Department of Communication
Texas A&M University
4234 TAMU
College Station, TX 77843-4234

Email Address: iahmed@tamu.edu

Education: Ph.D., Communication, Texas A&M University, 2009
M.A., Communication, West Texas A&M University, 2004
M.A., Mass Communication and Journalism, Dhaka University, 1997