Information Sharing

in

Distance Collaboration:

A Software Engineering Perspective

by

Laura Challman Anderson, BA; MLS; MS

Submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

School of Information Systems

Science and Engineering Faculty

Queensland University of Technology

2016

Supervisory Panel

San Jose Supervisor
Professor Geoffrey Z. Liu, PhD
School of Information
San Jose State University

Principal Supervisor
Professor Christine Bruce, PhD
School of Information Systems
Queensland University of Technology

Associate Supervisor
Professor Sharmistha Dey, PhD
School of Information Systems
Queensland University of Technology

Keywords

Information sharing, distance collaboration, online collaboration, virtual collaboration, activity theory, repertory grid elicitation, software engineering workgroups, agile software development, Information Sharing Distance Model, Information Sharing Discrete Distance Metric.

Abstract

The world seems smaller than ever today when we consider the global linkages enabled by technology that are possible between geographically separated people. Communication and information technologies to augment our human capabilities that were only in their infancy 50 years ago are everyday tools for communication and information sharing in our personal and professional lives. However, all of these dispersed people and tools give rise to a question about what people actually *d*o when sharing information in order to accomplish their work and what factors are important. Software developers, in particular, often work in heterogeneous workgroups, through distance collaboration, and using agile processes. This type of workgroup composition frequently includes geographic and time zone distance, attributable to industry developments such as globalization, open source projects, and outsourcing. They can also include cross-organizational, multi-disciplinary, heterogeneous roles (e.g., developer, manager), and workgroups with people joining mid-project. The co-occurrence of these dimensions of distance can create even more pressure for software engineering projects already challenged by an inadequate rate of success. Three dimensions of distance (geographic, time zone and multi-disciplinary) in information sharing and collaboration have been studied extensively, and cross-organizational information sharing to a lesser extent. But distance due to varying project tenure and heterogeneous roles have had much less research focus, and there have been few studies of information sharing across multiple distance dimensions.

The focus of this study of a Fortune 500 company workgroup is the examination of information sharing across six core distance dimensions in workgroups: (1) geographic, (2) time zone, (3) organizational, (4) multi-discipline, (5) heterogeneous roles, and (6) varying project tenure. This study extends the consideration of information sharing in distance collaboration beyond geography, time zone, organization and discipline to consider role and project tenure. It considers the implication of information sharing across all six dimensions of distance through a longitudinal 16-month study of a corporate software engineering project through an activity theory lens. The study finds that:

(a) Individual dimensions of distance, and especially multi-dimensional distance affects both information sharing and collaboration,
(b) A loosely coupled configuration can effectively handle geographical and time zone distance, rather than a tightly coupled approach,
(c) Sharing of information through mechanisms that are closest to the direct work of the sharer occurs frequently,
(d) Gaps created by discipline, organizational, and role distance are often unacknowledged and unaccommodated, while geographic and time zone gaps are explicitly addressed, but hampered by often inadequate technology solutions, and
(e) The introduction of new terminology occurs implicitly and incrementally in work sessions, not explicitly defined, but spread through usage in the collaborative activities.

This research proposes an Information Sharing Distance Model, which provides a framework for categorizing different factors in information sharing with six dimensions of distance at the core, as well as an Information Sharing Discrete Distance Metric, which represents the cumulative distance in a workgroup between collaborators across the six core dimensions of distance. This study also makes a method contribution through the demonstrated use of repertory grid elicitation with activity theory.

This research suggests that the effectiveness of workgroups with distance may be improved through mindful conducts of information sharing: proactive consideration of, and explicit adjustment for, the distances of the recipient when sharing information, to improve understanding and collaboration.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACM	Association for Computing Machinery
CIB	Collaborative Information Behavior
CIS&R	Collaborative Information Seeking and Retrieval
CSCL	Computer-Supported Collaborative Learning
CSCW	Computer-Supported Cooperative Work
ELIS	Everyday life Information Seeking
HCI	Human Computer Interaction
ISCAR	International Society for Cultural and Activity Theory
ISP	Information Search Process
LIS	Library and Information Science
PDA	Personal Digital Assistant
QA	Quality Assurance (process in Software Engineering)
RGT	Repertory Grid Technique
ZPD	Zone of Proximal Development

Definitions

Agile Software Engineering: An alternative software development method to plan-based, traditional approach to development of software. The 2001 Agile Manifesto (Holmström, Fitzgerald, Agerfalk & Conchúir, 2006) proposed four core values for software development and software engineering:

1. Individuals and interactions over processes and tools.
2. Working software over comprehensive documentation.
3. Customer collaboration over contract negotiation.
4. Responding to change over following a plan.
   (Dybå & Dingsøyr, 2008, p. 835).

Artifact: (noun) "An object made or modified by human workmanship, as opposed to one formed by natural processes." (Oxford English Dictionary Online, 2015).

Collaboration: (noun) "United labour, co-operation; esp. in literary, artistic, or scientific work." (Oxford English Dictionary Online, 2011).

Collaborative Information Behavior: "An umbrella term to connote the collaborative aspects of information seeking, retrieval, and use. We define CIB as the totality of behavior exhibited when people work together to (a) understand and formulate an information need through the help of shared representations; (b) seek the needed information through a cyclical process of searching, retrieving, and sharing; and (c) put the found information to use." (Karunakaran, Reddy & Spence, 2013, p.2438)

Common ground:

(1) "The knowledge, beliefs, and suppositions [that participants in a joint activity] believe they share about the activity." (Clark, 1996, p.38)

(2) "Mutual knowledge, mutual beliefs, and mutual assumptions." (Clark & Brennan, 1991, p. 127)

Information:

(1) "Anything of importance in answering a question...Information is what can answer important questions related to the activities of the target group." (Capurro & Hjørland, 2003, p. 390).
(2) "A difference which makes a difference." (Bateson, 1972, p. 463)

Information sharing: "An umbrella concept that covers a wide range of collaboration behaviors from sharing accidentally encountered information to collaborative query formulation and retrieval." (Talja, 2002, p.145).

Knowledge: "A set of symbols that represent thoughts, which the individual justifiably believes that they are true. In this analysis, information is a type of knowledge. It is neither an intermediate stage between data and knowledge, nor a synonym for knowledge." (Zins, 2006, p.459)

Software: "A program and all of the associated information and materials needed to support its installation, operation, repair, and enhancement." (Humphrey, 1989, p. 82).

Software engineering: "Refers to the disciplined application of engineering, scientific, and mathematical principles and methods to the economical production of quality." (Humphrey, 1989, p. 82).

Software requirements: "For a software project, requirements are specified in terms of software functionality, features, non-functional requirements of accuracy, speed, scale, reliability, maintainability, etc." (Agarwal & Rathod, 2006, p. 361).

Virtual teams: "Groups of geographically, organizationally, and/or time dispersed workers brought together by information and telecommunication technologies to accomplish one or more organizational tasks." (Powell, Piccoli & Ives, 2004, p. 7).

Activity theory definitions

Artifact/Tools: "The mediating artifact/tool can include artifacts, social others, and prior knowledge that contribute to the subject's mediated action experiences within the activity." (Yamagata-Lynch, 2010, p.16)

Knotworking: "A boundary-crossing, collective way of organizing work." (Engeström, Kaatrakoski, Kaiponen, Lahikainen, Laitinen, Myllys, Rantavuori & Sinikara, 2012, p.388)

Mediation: "…activity theory is specifically concerned with tools as means that mediate activity as a whole, rather than signs, that is, means that mediate specific mental operations…Tool mediation allows for appropriating socially developed forms of acting in the world." (Kaptelinin, Nardi, & Carroll, 2012, p.31)

Object: "The object is the goal of the activity." (Yamagata-Lynch, 2010, p.16)

Subject: "The subject … is the individual or individuals engaged in the activity." (Yamagata-Lynch, 2010, p.16)

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

QUT Verified Signature

Signature:

Date: December 24, 2015

Acknowledgements

This has been a journey marked and enriched by the company of companions, and I am deeply grateful to each person who has walked this road with me over these past six years.

First thanks go to my family: my husband and life partner Thomas, our daughter Caitlin and son David, and my mom Florence; and to the extended Challman and Anderson families – thank you for your unfailing support and encouragement. I also remember my dad Ron (1929-2004) with gratitude.

I am deeply indebted to my supervisory team: Dr. Christine Bruce, Dr. Geoffrey Liu, Dr. Helen Partridge, and Dr. Sharmistha Dey for their wise counsel and collaboration on this research in small to large ways. I am very grateful to Dr. Kristen Rebmann for her careful reading of the thesis and her many helpful suggestions. Thanks, too, to the study participants, whose kindness in talking with me and sharing their experiences enabled this research.

The IBM Research – Almaden community (and alumni) is a huge source of inspiration, ideas and support: a thank you for management support for my PhD program goes to Dr.Laura Haas, Marcia Bencala, Dr. Jeff Welser, Dr. Jim Spohrer, Dr. Stefan Nusser, and Dr. Ruoyi Zhou. Thanks, too, to my colleagues Dr. Cheryl Kieliszewski, Dr. Susan Stucky, Jeannette Iovino, Dr. Obinna Anya, Dr. Julia Rice and Ignacio Terrizzino.

It is also important to mention the faculty members and PhD students from the San Jose Gateway PhD program; it is a tremendous gift to call you colleagues and friends, and I appreciate being a member of this vibrant research community. Special thanks go to members of my cohort Dr. Clarence Maybee, Dr. Cheryl Stenström, and Dr. Diana Thormoto; to my dear mentor Dr. Virginia Tucker, and to Dr. Tina Inzerilla, Dr. Sylvia Edwards (QUT), Dr. Sandy Hirsh, Dr. Judy Weedman, Dr. Bill Fisher, and Dr. Mary Ann Harlan.

Finally, I am very grateful to my friends Dr. Dineh Davis, Resa Pearson, and Alexina Brooks for their support and friendship over many years.

> “Somewhere we know that without silence words lose their meaning,
> that without listening speaking no longer heals,
> that without distance closeness cannot cure.”
>
> Henri Nouwen

Chapter 1: Introduction

1.1 APPROACHING THE STUDY OF INFORMATION SHARING

The world seems smaller than ever when we consider the technology-enabled linkages that are possible today between geographically separated people. Communication and information technologies to augment human capabilities that were only in their infancy 50 years ago (Engelbart & English, 1968) are everyday tools in our personal and professional lives. However, all of these dispersed people and tools give rise to a question about what people actually *do* when sharing information using these technologies in order to get their work done. That is the motivation of this study in a single, specific information sharing setting.

The dynamics are complex in workgroups such as a geographically dispersed software engineering collaboration team in an industrial research setting. It is unrealistic (and impossible) to isolate any single phenomenon that may be in play across multiple people, world-wide labs, multiple time zones, varied roles, and multiple company units. It *is* possible, however, to examine the dynamics of a complex activity using a capable theoretical framework to see how the phenomenon exists and unfolds in that context. This study examines information sharing in a corporate software engineering activity, through an activity theory lens. The overall research aim for this study was to gain insight on what people do when they share information, how they think about sharing, what they share, what others share with them, and what this all means for collaboration. This chapter provides an overview of the thesis, the business context, the framing of the current literature for the context, the theoretical framework, and the specifics of the study, which include the research problem, questions, design, and contributions. It ends with summaries of remaining chapters.

The front matter contains general definitions, as well as definitions specific to activity theory, the theoretical framework of this study.

1.2 CONTRADICTORY FORCES IN SOFTWARE ENGINEERING TODAY

In the recent past, many businesses and organizations -- particularly in information technology -- have transitioned to a globalized operational model, moving work to people across geographical boundaries (Gilson, Maynard, Young, Vartiainen & Hakonen, 2015; Powell et al., 2004) utilizing information and communications technologies. It is routine today for software engineers to collaborate regularly over the Internet or via corporate networks, and to work as members of a distributed team across disparate time zones (Bjørn, Esbensen, Jensen & Matthiesen, 2014; Dalal & Chhillar, 2012; Herbsleb, 2007). While there are many advantages to distributed software engineering, the increased complexity brings added risks and challenges, most notably in the areas of collaboration and sharing of information related to a project (Koppman & Gupta, 2014; Cramton, 2001). The industry has also seen examples lately of companies pulling employees back into a centralized office configuration with the intent of improving collaboration (Oldham & Da Silva, 2015; Goudreau, 2013; Swisher, 2013). This begs the question of how to incorporate people who, due to physical distance or other reasons, cannot work in a centralized office location.

The trend of dispersed employees complicates an already problematic reality of low software project success rates. Success, in the most general sense, refers to the creation of software that meets customer requirements, in the agreed upon timeframe and budget, and which can be put to productive use.[1] There is a broad agreement both in academia (Lehtinen, Mäntylä, Vanhanen, Itkonen & Lassenius, 2014; Agarwal & Rathod, 2006) and among software industry practitioners about the low success rate of software engineering projects, by almost any measure. Multiple industry studies have shown that only 17-50% of software projects were considered to be "successful", even when project cost overruns and software functionality reductions are excluded from consideration. Inclusion of cost overruns and software scope reduction would reduce the success rate to an even

[1] Defining success for software projects is outside the scope of this study, but inhibitors to success include customer dissatisfaction, inability to put the developed software into production, and failure to realize expected business value, among other factors.

lower number. Beginning in 1994, The Standish Group (2013) began an annual IT industry analysis, revealing extremely low software project success rates in their ongoing reports series. They showed that the overall industry software project success rate was only 16% in 1994, improving to a modest 33% in 2013 (Ambler, 2014). The Standish Group's findings were independently confirmed in a 2004 report published by the United Kingdom Royal Academy of Engineers and the British Computer Society (Hussey, 2005). The 2004 document reported a meager 16% project success rate as well as economic loss of billions of pounds per year on these projects throughout the European Union. Recent literature continues focusing on the problem of software project productivity (Jorgensen, 2014; Jangir, Gupta & Agrawal, 2012).

The characterization of a crisis in software engineering projects has been disputed (Ambler, 2014) based on an independent survey of software projects showing success rates ranging from 49% to 72%. However, even with these higher rates of success, there is still considerable room for improvement and reduction of wasted resources.

The software industry and the research community have struggled to uncover factors contributing to the general problem of low success rates of software engineering projects, in hopes of improvement. Some people feel that the agile software engineering and project management methodologies developed and implemented widely in recent years can improve the effectiveness of software engineering projects (Beecham, Noll & Richardson, 2014). While agile methodologies do show many benefits, the tightly coupled work style and the lack of written documentation can be especially challenging for geographically dispersed teams. The confluence of geographically dispersed workgroups and agile methodologies, while not incompatible, are at odds.

Information sharing is integral, and critical, to project collaboration, and even more so for distance collaboration. The processes, mechanisms, artifacts, and avenues of information sharing in distance collaboration are important to examine, considering not only geographic distance and time zone, but also other dimensions that can separate people, such as disciplinary background, organization and company boundaries, different job roles in the work activity, and length of tenure on the project. Even though research on information sharing has increased over the past ten years, research on information sharing in working software development teams is scarce. Recent changes, most notably the rise of heterogeneous workgroups, new technologies, and significant changes in work environment configuration, suggest that additional studies in this space are needed.

This study addresses that gap through an examination of the information sharing mechanisms in a complex corporate software engineering activity, with a focus on these important core dimensions of distance in a workgroup.

1.2.1 Agile software engineering

According to Humphrey (1989), software is "a program and all of the associated information and materials needed to support its installation, operation, repair, and enhancement", and software engineering "refers to the disciplined application of engineering, scientific, and mathematical principles and methods to the economical production of quality" (p.82). The development team of a software system in a corporate setting can range from a few people, to a workgroup of dozens of people, to very large workgroups that number in the hundreds. While the complexity increases with the size of workgroups, even a workgroup of two people must coordinate, share information, and collaborate in this intellectual, knowledge-intensive work.

Software engineering projects are problematic from at least four perspectives: they have a low industry success rate; they are a high-cost undertaking; they are complex and complicated to manage; and overall they often underperform in meeting important project objectives (Agarwal & Rathod, 2006; Linberg, 1999; May, 1998). Furthermore, the recent rise of virtual teams that are composed of members distributed globally has only increased the challenges and risks to effective collaboration (Herbsleb, 2007; Carmel & Agarwal, 2001) by adding additional obstacles, such as differing time zones and languages.

Agile software engineering is a newer approach that has become popular in the past few years in an attempt to reduce project risks, particularly software delivery delays and failure to meet user requirements. The heterogeneous composition of an agile software engineering team can include the end users, the customers, and the extended stakeholders, in addition to the design, architecture, and implementation team (project managers and software engineers). In this study, the project workgroup (and participant scope) included the core implementation workgroup of technical project leaders, the researcher/software engineers, as well as extended project workgroup members and stakeholders.

There is a diverse mix of direct and indirect participants in the process of agile software engineering: the software engineers implementing the software, customers reviewing prototypes and giving feedback and suggestions, and executive management monitoring the project periodically. Expansion of the project workgroup to include this broad array of people has many benefits, but also can inject some issues. In particular, people from varying disciplinary backgrounds may bring diverse perspectives, terminologies, and embedded knowledge to the task at hand. These differences can create communication challenges and the need to invest time in establishing common ground, but may also result in a better project outcome.

Agile software engineering is only one class of a variety of methodologies by which software can be developed. Figure 1-1 illustrates a basic waterfall sequence of events in sequential software development, and Table 1-1 details and contrasts the characteristics of the two models. Looking at Figure 1-1, the sequence of this sequential method is visible, colloquially called "waterfall" since one step flows over/down to the next step, and so on, until the conclusion of the project. It is a traditional software engineering method built on engineering processes. It includes creation of detailed documents at the beginning of each phase, and methodical specification in advance of the software to be developed. While this disciplined process has many benefits, it is typified by misunderstandings among the technical workgroups and the users about the functionality of the software developed, often accompanied by delays and schedule overruns.

The agile family of methodologies arose to address the shortcomings of the traditional waterfall method through close collaboration with the users and frequent integration of working code for user review. .Boehm's (2002) table (reproduced as Table 1-1 above) contrasts agile and plan-driven methods in seven key facets of software. Agile methods embody a flexible and dynamic approach, emphasizing a collocated and collaborative environment, with emergent requirements and changes, and focus on rapidly realizing value for the customer. This contrasts

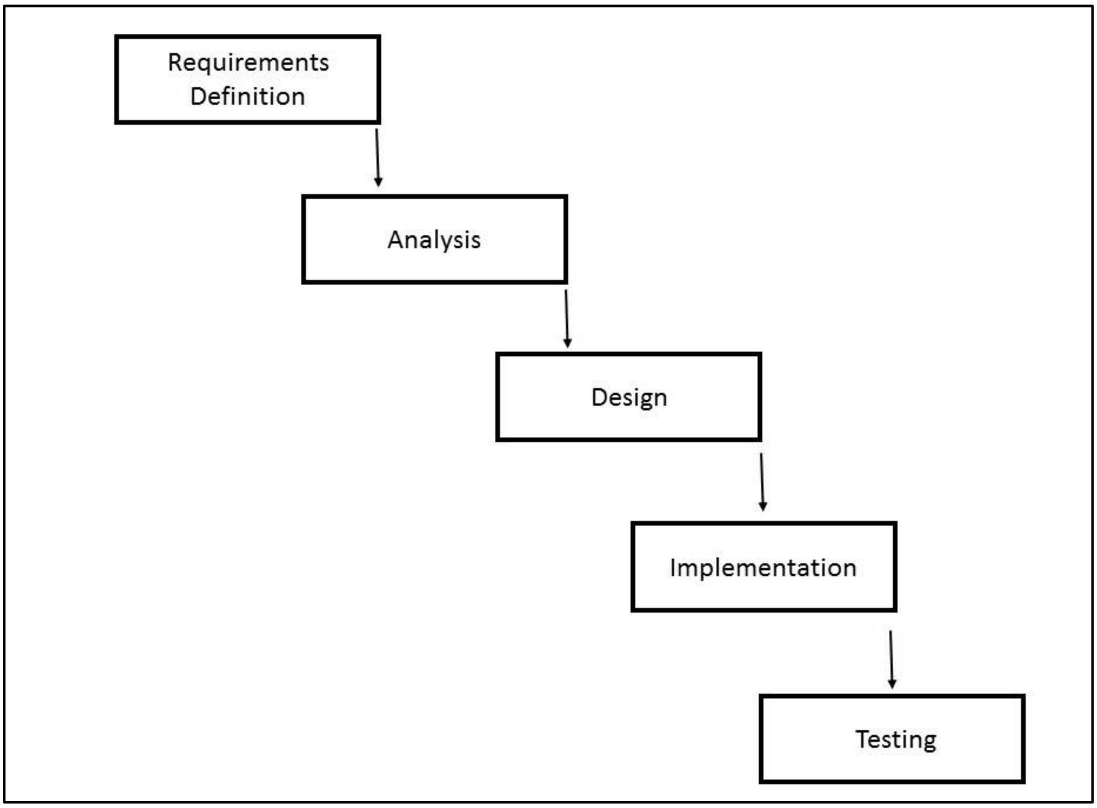


Figure 1-1: Waterfall software engineering method (adapted from Bassil, 2012, p.743)

Home-ground area	Agile methods	Plan-driven methods
Developers	Agile, knowledgeable, collocated, collaborative	Plan-oriented; adequate skills; access to external
Customers	Dedicated, knowledgeable, collocated, collaborative, representative, empowered	Access to knowledgeable, collaborative, representative and empowered customers
Requirements	Largely emergent, rapid change	Knowable early; largely stable
Architecture	Designed for current requirements	Designed for current and foreseeable requirements
Refactoring	Inexpensive	Expensive
Size	Smaller teams and products	Larger teams and products
Primary objective	Rapid value	High assurance

Table 1-1: Characteristics of agile and plan-driven methods (Boehm 2002, p.68)

with traditional waterfall or plan-driven methods that are typically more expensive overall, are comprised of larger teams and products, invest more time and resources at the beginning of a project, and have a more well-understood and stable set of requirements. The flexibility of a changing project and the lack of written documentation combine to create challenges in information sharing across the agile project team. Literature about these software engineering methodologies are covered in more detail in Chapter 2.

1.3 PHENOMENA IN FOCUS IN THE STUDY

This section frames the study focus: the foreground topic of information sharing, and the background topic of distance collaboration, and discusses each in detail.

1.3.1 Information sharing

The primary activity of interest in this study is information sharing during the lifespan of a project -- a collaborative, knowledge-intensive, and complex setting. Previous studies of information sharing have focused on a variety of settings and participants, such as students in academic settings, group learning contexts, homogeneous groups of people (e.g., engineers), healthcare settings, and emergency response settings, to name just a few (Pilerot, 2014; Grubb & Begel, 2012; Mesmer-Magnus, DeChurch, Jimenez-Rodriguez, Wildman & Shuffler, 2011; Allard, Levine & Tenopir, 2009; Golovchinsky, Qvarfordt & Pickens, 2009; Cho, 2008; Evans & Chi, 2008; Fidel, Bruce, Pejtersen, Dumais, Grudin & Poltrock, 2000). Increased knowledge and understanding about information sharing in these contexts has the potential to address and solve many problems related to collaboration (Poltrock, Grudin, Dumais, Fidel, Bruce, Pejtersen, 2003). There is significant research on engineers' information sharing in a collaborative context, particularly in corporate environments, but studies of multidimensional distance in a workgroup, including interdisciplinarity, time on project, role, and matrixed organizational configurations, are few. Few studies could be found (outside of healthcare) of an intact, multi-disciplinary, and multi-role workgroup with a focus on information sharing (Nissen, Evald & Clarke, 2014; Talja & Hansen, 2006; Hertzum & Pejtersen, 2000) with a unit of analysis at the activity level.

Information sharing (also referred to as "knowledge sharing" in some literature) is identified as a component of "information use" within the field of information seeking research, but only as a sub-concept and with emphasis on the receiving end. This makes it confusing to talk about the proactive activity of "sharing", because it is conceptually situated within a larger overall concept of information seeking and it is the flipside of "seeking" activity. For example, Talja (2002) frames the work of information sharing within the topic of information seeking while developing a framework for fine-grained types of information sharing in the academic community: " 'Information sharing' is used as an umbrella concept that covers a wide range of collaboration behaviors from sharing accidentally encountered information to collaborative query formulation and retrieval." (Talja, 2002, p.145)

The term "information sharing" is used across the disciplines to discuss specific configurations of human information sharing, some crossing over to computer and agent-based systems and modeling. Information sharing is seen in formal information organizing systems such as Management Information Systems (Barrett & Konsynski, 1982); database systems (Agrawal, Evfimievski & Srikant, 2003); emergency response systems (Aedo, Diaz, Carroll, Convertino & Rosson, 2009); in market models in the field of economics (Gal-Or, 1995); for purposes of online social networking (Acquisti & Gross, 2006); and in multi-agent systems (Foner, 1995).

In the information seeking literature, a complementary view to Talja's (2002) characterization of information sharing is that it is "the act of providing a helpful answer to a request for information" (Rafaeli & Raban, 2005, p. 63). In this study, information sharing is defined as experiences and practices of information dissemination and communication, ranging from spontaneous sharing of information on known areas of mutual interest, uncertainty, or previously discussed topic, to the practice of approaching a colleague for answer to a question which is related to the common project tasks and objectives.

A related term, "collaborative information seeking", is also used in recent literature to reference an open population collaborating voluntarily in online community settings (such as forums and blogs) to address their (individual) information needs (Foley & Smeaton, 2009; Golovchinsky et al., 2009; Evans &

Chi, 2008). However, the term has also been used to describe workgroup members' collective and individual behaviors of information seeking while working on a collaborative project. Hertzum (2008) defines collaborative information seeking as "the information-seeking activities performed by actors to inform their collaborative work combined with the collaborative-grounding activities involved in making this information part of the actors' shared understanding of their work" (p. 958).

Collaborative behaviors of information seeking in the context of software engineering range from working together to find an answer to an information gap, to consulting with an expert colleague to solve a problem, and to conveying the results of a work item that has cascading implications in the project. Broadly, these include seeking, searching, retrieving, using, and sharing. It is information sharing in the specific context of a software engineering project that this research is intended to investigate. The central concept of "information sharing" in this research is defined and interpreted more in line of Hertzum's (2008) definition of collaborative information seeking in the previous paragraph..

1.3.2 Distance collaboration

Collaboration is an important topic since many believe it to be a critical success factor for the workgroup in a wide variety of settings, including technology, the service industry, academia, and healthcare. There is a wide body of literature examining the many facets of collaboration across these multiple contexts. Multiple authors point out the imprecise use of the terms "collaboration" and "teamwork" almost interchangeably, when in fact they can be two very different phenomena (Croker, Higgs & Trede, 2009; Cummings & Kiesler, 2008). There can certainly be some work being done within a team (teamwork) without true and effective collaboration, to name just one example.

Geography is the classic distance dimension in collaboration. Often the term "distance" is used to mean geographical distance, and it is the focus of a large body of scholarly literature. Geographically-distributed workgroups may face some intrinsic issues due to the lack of physical proximity and a set of cascading issues that can arise as a result. In response, they often form cross-organizational reporting structures spanning formal boundaries. This configuration provides flexibility and responsiveness for a project, but also multiplies the number of additional stakeholders and management personnel with a stake in the project, adding complexity. A variant of widely dispersed teams, called *near-shoring* (Carmel & Abbott, 2007) —where teams are separated geographically but share a similar time zone— is another model to provide increased overlap of working hours in the normal workday. The additional distance dimensions of project tenure, role, discipline, and organization are also important in information sharing and collaboration, and broaden the consideration of distance from simple proximity to more nuanced distances.

1.4 INFORMATION SHARING AND CORE DIMENSIONS OF DISTANCE

This research focuses a lens on how, why, and when information sharing occurs in a particular collaborative work context. Stated another way, this research examines when the action "to inform" another person occurs in a specific, complex setting from the perspective of the people. It examines both the sender and receiver of the shared information for a view of what such action means to other aspects of the collaborative activities and what effects additional dimensions of distance may have on information sharing.

This study explores and investigates information sharing between people collaborating on a software engineering project using an agile development approach across diverse dimensions of distance (graphically depicted in Figure 1-2.) There are also many other factors that have been studied in information sharing and collaboration, such as trust, power, and preferences (Yang & Maxwell, 2011), to name only a few. However, the focus and scope of this study is to investigate distance factors in information sharing -- across geography, time zone, discipline, tenure on project, role, and organization, with particular attention to the under-researched implications of project team composition of a range of formal roles (Kamal, Weerakkody & Irani, 2011; van Knippenberg & Schippers, 2007) and differing project tenure (Gilson, Lim, Luciano & Choi, 2013). Stated very simply, the study looks at the activities, mechanisms, and artifacts intended "to inform" and also explores what information sharing means to collaborators in the project.

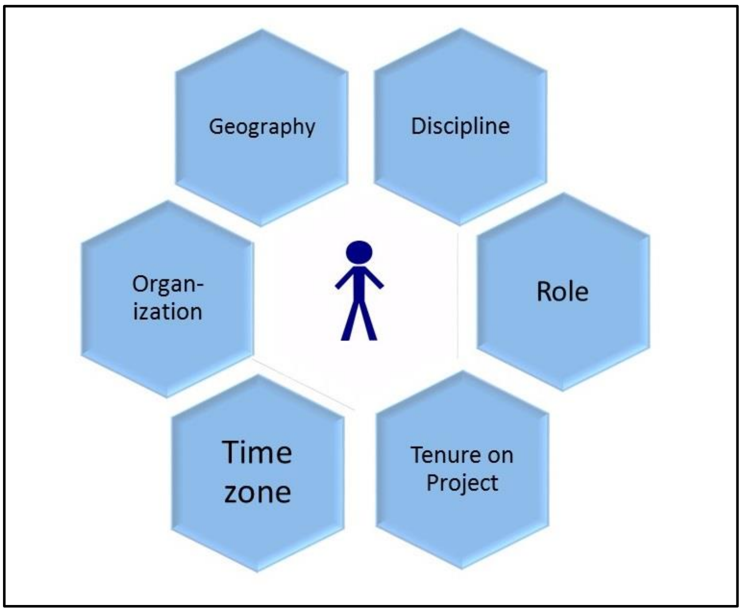


Figure 1-2: Dimensions of distance

1.5 RESEARCH QUESTIONS

The core question of this research is: “How does information sharing occur in the distance collaboration of virtual teams?” The research problem centers on the phenomenon of information sharing in workgroups, and the associated challenges to collaboration and the accomplishment of work objectives by an extended project team. The subordinate research questions for the study are:

1. How do information sharing activities manifest themselves in distance collaboration?
2. When and in what kinds of circumstances does information sharing occur in distance collaboration?
3. What types of information sharing behaviors and forms of shared information can be identified?
4. What attributes are related to different types and forms of information sharing in distance collaboration?
5. What purposes does information sharing serve in distance collaboration?

These questions were adapted from the work of Talja and Hansen (2006).

1.6 AN ACTIVITY THEORY APPROACH

The theoretical framework for this study is activity theory (Engeström, 1987; Leont’v, 1978; Vygotsky, 1978), sometimes called cultural-historical activity theory. Activity theory grew out of a Soviet psychological school of thought called cultural-historical psychology (Vygotsky, 1978) early in the 1900s. Marxism influenced activity theory, particularly the labor aspect. Although developed many years prior, Vygotsky's work first came to the Western world in the 1970s. Since that time, many disciplines (e.g., human-computer interaction (HCI), management consulting and healthcare) have embraced activity theory to provide analytical insight. Leont'v (1974), Engeström (1987), Cole (1999), and others[2] have continued to develop activity, expanding it in several directions. With a focus on human motivation, human activity, and the externalization of human ideas and objectives into something tangible in the world, it provides the capability for a systemic visualization and abstraction of complex systems. Activity theory provides a framework for understanding the interactions and interplay among people (subjects), the human objective of the activity (object), the tools and mediating artifact, and the outcome in a larger social, historical, or work context (Widén-Wulff & Davenport, 2007). It simultaneously enables analysis at the level of the individual and larger collections of people such as an organization or community. As a theoretical construct for this study, the term activity theory refers to the Vygotsky-Engeström line of cultural-historical activity theory.

The activity system framework provides a way to characterize the situating context of the teamwork, along with participant viewpoints and the structural components of the collaborative environment. The situated context of the workgroup is framed in the activity theory system, and analyzed within the structures of Subjects, Objects, Instruments (Tools), Community, Rules, Division of Labor, and Outcome. This provides an abstraction to think about the structural components of the environment, the characteristics and intent of each participant, and participants' perspectives on the overall project. Activity theory provides an insightful framework to look at the systemic contradictions and tensions that

[2] There are two "adjacent" research threads that have their roots in the work of Vygotsky's work on cultural context and mediation, but are distinct from activity theory. These are: (1) sociocultural theory of mediated action (Wertsch, del Rio, & Alvarez, 1995); and (2) the theory of situated learning (Lave & Wenger, 1991; Suchman, 1987). Cultural-historical activity theory is more comprehensive than sociocultural theory in including collective action. The term "sociocultural" has also been used more generally as a term for Cultural-Historical Activity Theory (Lantolf & Thorne, 2007), creating a confusing situation. Activity theory is better suited to incorporate unanticipated aspects of growth and change than the theory of situated learning (Engeström., Miettinen, & Punamäki-Gitai, 1999, pp. 11-12).

arise within and between the components of the activity system. An example of this is the uniquely situated view of each person (subject) in an activity system and his/her individualized view of the desired outcome of the activity.

This study utilizes activity theory because of its strength in including the wide and complex context, the multiple human participants, and both the expected and unexpected developments. It is powerful in providing a complex environment for comprehensive data gathering and deep analysis of the phenomenon of information sharing. Finally, activity theory's mediation construct (tools, artifacts) contributes a unique focus for examining mediation in information sharing.

In this research, the study participants are associated with a single project, providing the opportunity for an end-to-end look at one context (software engineering) from multiple vantage points (extended project workgroup and stakeholders).

This research studied the case of an industrial software engineering workgroup featuring the following characteristics:

1. Physically dispersed: The workgroup was globally distributed across geography and time zones;
2. Cross-disciplinary: It included professionals trained in computer science, social science, science, and business, and project management;
3. Heterogeneous in roles: Workgroup members included software developers, technical leaders, architects, users, project managers, and people managers;
4. Intensive in information sharing: Workgroup members frequently shared information related to the software engineering process.

The implementation of this study began after the project started, but early in the lifecycle, with data gathering conducted through semi-structured interviews utilizing the repertory grid elicitation technique. The interviews were audio recorded and transcribed, and analyzed using activity system analysis and the Leximancer 4 content analysis software, using the constructs of activity theory. This approach enabled both an activity system perspective as well as conceptual and thematic analysis

1.7 CONTRIBUTIONS OF THE RESEARCH

Developed in this study is a new model of the cumulative information sharing distance between collaborators (Figure 4-3) called the Information Sharing Distance Metric. This model illustrates the core aspects of heterogeneity, experienced as distance by the members of a workgroup, which accumulate to form a cumulative distance between collaborators, over which shared information must traverse. The model integrates disparate factors from previous studies into a cohesive core framework of factual distance dimensions, which include geography, time zones, organizational distance and multi-discipline, project role, and project tenure. Studies of two of these dimensions - role and project tenure - are fewer in number than other aspects, and important to understand, because heterogeneous roles in a project and workgroup member changes during a project are frequent occurrences in industry.

Additional original contributions to the existing body of knowledge from this research include:

- Increased understanding of information sharing in a heterogeneous workgroup (not just engineers) across disciplinary boundaries;
- Insights about the meaning and usage of shared information artifacts when utilized by people from different organizations or with different roles, and about how they can serve a transformative purpose; and
- The application of the activity theory theoretical framework to this highly specific setting, augmenting the research studies at the Center for Research on Activity, Development and Learning at the University of Helsinki.

Although the general information behavior of engineers has been extensively studied, and specific influencing factors, often single ones, researched across multiple populations, the focus on the distance factor, and framing of dimensions of distance in a core configuration in this study bridges a gap in the research literature. This study is significant, and addresses the gap, of a primary focus on information sharing across multiple dimensions of distance (or stated another way, across multiple factors of heterogeneity) along with a secondary perspective on collaboration in an agile software engineering context. This study also identifies a framework and a set of core dimensions of distance for exploration in the future to understand information sharing and collaboration more deeply, in the theoretical dimension as well as in the pragmatic setting of software project management. Findings from this stream of research may suggest practical changes to improve both information sharing and collaboration.

1.8 CHAPTER ABSTRACTS

1.8.1 Chapter 2 - Literature Review

The second chapter provides a general overview of the foundational literature in the information sharing landscape and how information sharing emerged from information seeking within the information behavior area. The cross-disciplinary treatment outside of the field of Library and Information Science is detailed, and how previous scholars characterize and understand information sharing as well. The theories applied to information sharing and findings from previous empirical studies are also summarized. The perspective on information sharing adopted in this study is detailed, along with identification of the gaps in the literature around information sharing over distance. Next is selective coverage of the background topics of distance collaboration and agile software engineering, as a situated backdrop for the primary focus of study. Finally, the literature supporting the theoretical framework, activity theory, is reviewed.

1.8.2 Chapter 3 - Research Design

This chapter describes the overall research design, with a focus on the theoretical lens and associated methods. First, the overall research philosophy and approach are presented, followed by a discussion of the core research question, the general problem space for the study, and the subordinate research questions to be explored. Following that are several sections on activity theory, the primary theoretical framework, starting with the application of activity theory. A discussion of the suitability of activity theory to answer the research questions of this study follows and an exposition of criticisms about activity theory in order to provide a comprehensive perspective on activity theory and on the epistemological discussions about it in the literature. Next is a discussion of other potential theories considered for application in this study, with an exposition of the rationale of deciding not to use them and explanation of why they are not considered the best choice for this study. In addition, this chapter presents the approach of data gathering and analysis, built on the repertory grid technique (Kelly, 1955; Adams-Webber, 2006), and a detailed description of the implementation of associated methods. This part also reports the two pilot studies done in advance, along with changes of the method of the main study because of learned lessons from the pilots. Next, the methods of the main study are detailed. The chapter concludes with the presentation of ethics clearance information.

1.8.3 Chapter 4 – Findings

Chapter 4 covers the findings from the data about people's information sharing and collaboration. Here we enter the individual and collective world of the participants and examine the information sharing mechanisms – both events and artifacts. Provided first is a single consolidated view of the information sharing mechanisms across the 23 unique contexts of the participants to give a sense of the landscape, followed by additional detail about information sharing activities, events, and artifacts identified in the study. Discussion continues with a review of activity systems of varying granularities, followed by a look at the longitudinal development of the highest-level iProject activity system over 19 months. Reviewed next are identified contradictions present at varying times, and innovations attributed to activity system changes in the "Zone of Proximal Development". Next is presentation of the ways knotworking, the dynamic and flexible relationships of multiple activity systems, emerges in the relationships between the peer projects of iProject. The final section contains observations and insights about collaboration gained from the data and a reprise of the overall findings in the form of answers to the research questions, as well as unexpected study findings.

1.8.4 Chapter 5 – Discussion

Chapter 5 contains a discussion of the key constructs of the study: activity theory, the research questions and the related findings. The focus here is first on confirmation of previous research, then what is new, followed by what is controversial. Next is the phenomenon of multidimensional distance, a critical aspect of information sharing, followed by how information sharing influences the collaboration between individuals and workgroups with distance dimension(s).

Discussion continues with the nature of the information sharing mechanisms, focusing on the two-way influence between information sharing and collaboration. Covered next are the model and metric for information sharing distance, developed in this study. Finally is a review of the innovative methodology of this study (activity theory with the repertory grid interview protocol, and Leximancer 4 data analysis).

1.8.5 Chapter 6 – Conclusion

Chapter 6 provides closure to the document, a summation of the key points of the dissertation work, a succinct statement of the answers to the research questions, and a recommendation for potential future work built on this research to further our collective understanding of information sharing and collaboration over distance.

Chapter 2: Literature Review

2.1 CHAPTER OVERVIEW

The study of information sharing (and knowledge sharing, an equivalent term) has increased over the past few years and across a variety of settings, especially in the disciplines of Library and Information Science (e.g., information behavior), Computing and Software Engineering (e.g., Computer Supported Cooperative Work, CSCW), and in the Business and Management disciplines (e.g., Knowledge Management). This chapter presents relevant literature from these disciplines, and identifies aspects of information sharing that have not been sufficiently studied yet, establishing a foundation for this study and identifying gaps to be addressed.

First, a general overview of the information sharing landscape is established: how it emerged from information seeking within the information behavior area, the cross-disciplinary treatment outside of the field of Library and Information Science, and how previous scholars characterize and understand information sharing. The theories applied to information sharing and findings from previous empirical studies are also presented. The perspective on information sharing adopted in this study is detailed, along with identification of the gaps in the literature around information sharing over collaboration distances (other than geographical distance). Next is selective coverage of the background topics of distance collaboration and agile software engineering, as a situated backdrop for the primary focus of study. Finally, the literature supporting the theoretical framework, activity theory, is reviewed.

The activity of information sharing is the foreground focus and the primary phenomenon of interest of this study, situated in a real-life, complex, and messy context featuring distance collaboration and agile software engineering. Figure 2-1 illustrates the interrelated topics in the literature review: the centrality of information sharing contextualized within distance collaboration and agile software engineering. There are many moving parts in play, and learning from the previous literature about these topics is helpful to frame the study, and to ensure a unique research

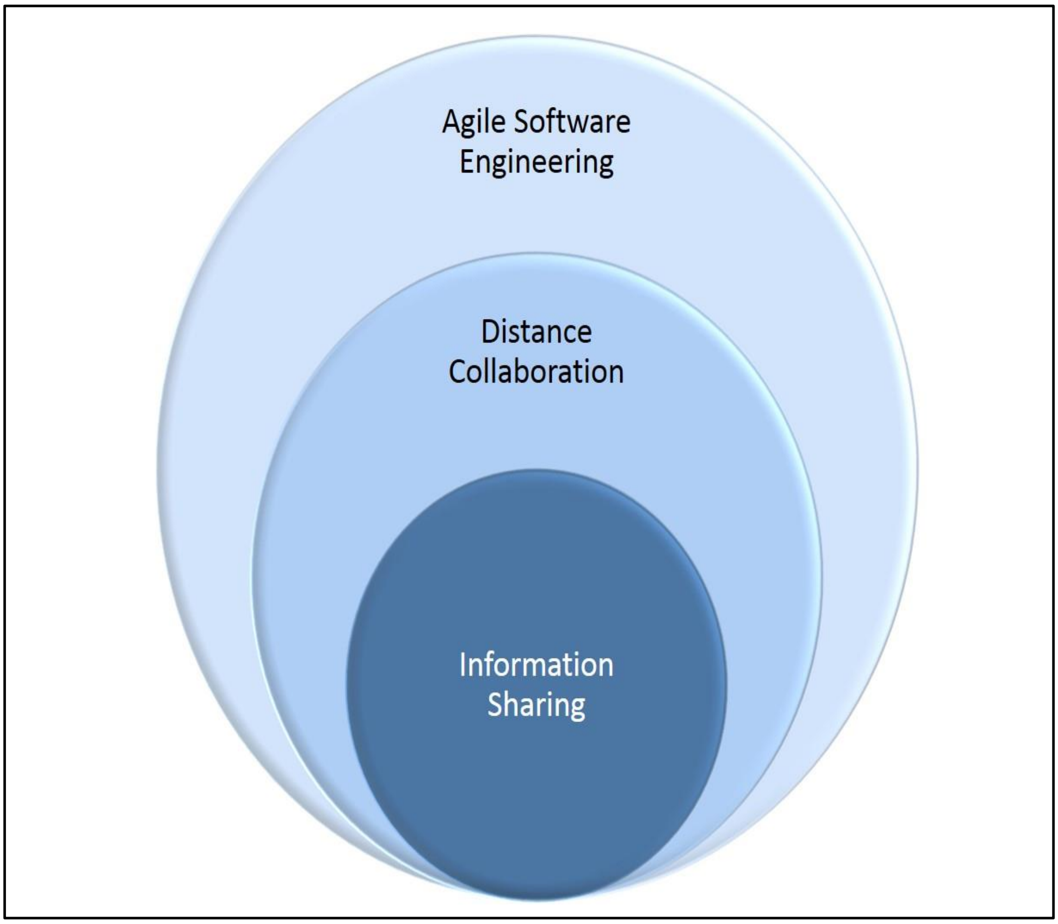


Figure 2-1: Visualization of foreground and background topics

contribution. These topical areas naturally situate the work in an inter-disciplinary space, due in large part to the capabilities that computing technologies and the Internet have enabled so ubiquitously, the importance of collaboration, and the relevance of information/knowledge sharing for so many disciplines.

Literature from these multiple disciplines is important, although literature from the Library and Information Science (LIS) field is a primary focus. The rationale for this approach is the rich body of research on information behavior, the centrality of information interaction and information experience within LIS, and the importance of information sharing as an information behavior. However, an examination of information sharing from these additional disciplinary perspectives provides a broader perspective on a human information behavior manifested in a wide variety of contexts.

2.2 INFORMATION SHARING: THE FOUNDATION

2.2.1 What is information sharing?

Information sharing (Wilson, 2010) has been characterized as a specific information behavior (Bates, 2010), and one that has been closely associated with the activity of information seeking (Case, 2012). It is an information activity within a spectrum of ways that human beings interact with and experience information. During the past 10-15 years, the research focus in information behavior studies has expanded from the information seeking and retrieval domain to a broader examination of aspects of other information behaviors (Wilson, 1999). Given the somewhat organic emergence of information sharing from this broad area, it is not surprising that there are multiple treatments of the term "information sharing". Definitions of information sharing range from high level, general characterizations of the activity, to the extremely specific:

> [Information sharing is] an umbrella concept that covers a wide range of collaboration behaviours from sharing accidentally encountered information to collaborative query formulation and retrieval. (Talja, 2002, p.145)
>
> … to provide information to others, either proactively or upon such that the information has an impact on another person's (or persons') image of the world, i.e., it changes the person's image of the world, and creates a shared, or mutually compatible working, understanding of the world. (Sonnenwald, 2006, p.1)
>
> Information sharing is the primary process through which teams utilize their available informational resources. (Mesmer-Magnus et al., 2011, p.215)
>
> Information sharing behaviour describes the explicit and implicit exchange and sharing of data between people, groups, organizations and technologies. (Widén & Hansen, 2012, p. 5 of 12)

The literature examines "information" itself from multiple vantage points, and it is useful to understand this treatment of information, since there are implications for "sharing", as characterized in Buckland (1991): (1) Information-as-process; (2) Information-as-knowledge; and (3) Information-as-thing (p. 351).

Looking at information in the sense of "information as thing" provides a tangible focus on data and documents that are informative, as Buckland (1991) noted. A lively exchange in the literature unfolded after his essay appeared, bringing to the field a healthy discussion about the external embodiment of information. Intangible cognitive processes are definitely involved in human information processing, but the ability to focus on a particular incarnation or representation of that which informs is helpful for empirical study. The materialization of information in the form of a document (Pilerot, 2014) is also a focus in the literature. Ongoing research needs to focus on smaller and more informal information forms, due to the rise of mobile technologies and text messaging -- where the shared information is more granular than a document.

Rafaeli and Raban (2005) defined the following three concepts in their review article on online information sharing, with an economic perspective:

> 1. '*Information'* is data that have been analysed and/or contextualized, carries a message and makes a difference as perceived by the receiver (Ahituv & Neumann, 1986).
> 2. *'Knowledge'* or 'expertise' is defined as a human quality that builds on data and information together with experience, values, and insight.
> 3. *'Information sharing'* is the act of providing a helpful answer to a request for Information. (Rafaeli & Raban, 2005, p.63)

Rafaeli and Raban's (2005) definitions of *information* and *information sharing* require the receiver to confirm the impact or value, in contrast to Buckland's (1991) treatment of the three types of information which is much broader. And their definition of information sharing limits this behavior to a response to a request and does not include unsolicited transfer to others, a modality common when people work together. Rafaeli and Rabin (2005) also made the economic observation that information and knowledge are "simultaneously private and public goods" (p. 64). This highlighted the individual and collective nature of information, and the dichotomy of both individual value and organizational value of information sharing in a work effort.

Talja and Hansen (2006) built on this idea of information providing an impact or difference, and quoted Reddy, Dourish and Pratt's (2001) observation in medical care setting studies that "...the information itself, in HealthStat, does not tell a complete story." (p.248). Talja and Hansen (2006) noted:

> Developing and sharing understandings of the information is as important as the availability of the information itself, and integrating pieces of information into the context of work often requires active collaboration (p. 126).

In the same time period (the past 10-15 years), there have also been many studies of either information sharing or knowledge sharing from other disciplines, most notably in computer science and business, with a focus on collaboration technologies and organizational knowledge management, respectively (Pilerot, 2012). Use of the terms "knowledge" and "information" are often interchangeable (Wang & Noe, 2010) and they noted, "Knowledge sharing refers to the provision of task information and know-how to help others" (p.117).

2.2.2 Working definitions for this study

"Knowledge sharing" and "information sharing are considered equivalent in this study, with the term "information sharing" used to represent both throughout the dissertation for consistency. Following are the working definitions of *information sharing* and *information sharing mechanisms*:

- Information sharing (v): An activity resulting in an information exchange between two or more people.
- Information sharing mechanism (n): A means to accomplish an information exchange between two or more people represented in a tangible artifact or an interactive event.

Information sharing activity is an umbrella term used to describe the general human activity in which information sharing occurs. "Information sharing mechanism" or "information sharing event" are more descriptive and granular terms used to describe information sharing at a more detailed level.

In this study, the distinction between data and information is not a critical one, as both are included, just as knowledge sharing and information sharing are both considered along a continuum. The focus in this study is on sharing between participating individuals to advance the work objective(s) of the project, and a distinction between knowledge and information is not made. Important is the fact that there is (or is not) an exchange, and the subsequent impact of the exchange in the workgroup, and in the project.

The approach taken in this study, described more fully in the Research Design Chapter (Chapter 3), builds on both Buckland's (1991) view of "information as thing" (p. 351) and Talja and Hansen's (2006) view about the criticality of developing shared understanding. The study focuses on the information sharing activity itself, the representations (artifacts and events) of information, and the meaning these impart to the collaborators.

2.2.3 Origin in multiple disciplines

Over the past 20 years or so, information sharing emerged from the area of information seeking research with a social manifestation and a strong cross-disciplinary base extending into the fields of Information Behavior (LIS), Computer-Supported Cooperative Work (Computer Science), and Knowledge Management (Organizational Management and Business).

In the Library and Information Science literature, information sharing is most closely associated with studies of information needs and information seeking (Taylor 1968), as a corollary activity of information use after obtaining information (Pilerot & Limberg, 2011; Wilson, 2010; Allard et al., 2009; Savolainen, 2009, 1993; Fisher & Julien, 2009; Algon, 1999). Kuhlthau's (1991) important work presented an information search process (ISP) model with six steps: initiation, selection, exploration, formulation, collection, and presentation. Although Kuhlthau's model did not directly address the sharing of raw information gathered from sources, the final stage of her ISP model – "presentation" – alluded to sharing of a new understanding or solution. In the larger context of information exchange and use, the new understanding and solution becomes "new" information being distributed, and users' sharing of a new understanding and solution constitutes one form of information sharing. As Kuhlthau noted:

> The ISP culminates in a new understanding or a solution which may be presented and shared. Evidence of the transformation of information into meaning is present in the products or presentations in which users share their new knowledge with others. (Kuhlthau, 1991, p. 361)

We see this early information behavior literature focused on individual information seeking or searching, but we also see that social dimensions unfold in the presentation stage. Demonstrated are information activities that often culminate in gaining new knowledge, or information subsequently shared with others. For that reason, information sharing emerged from information seeking and other information behaviors with a social manifestation.

Kuhlthau's (1991) work has two connections to the theoretical underpinnings of this research: first, to Kelly's (1955) personal construct theory, which supports the repertory grid technique used in this study, and her "zone of intervention" concept, based on Vygotsky's (1978) "zone of proximal development" from activity theory (Fourie, 2013). Chapter 3 (Research Design) contains a deeper discussion of the use of these components in this study. This shared foundation between important and established research in information seeking and this new research on information sharing demonstrates a shared lineage of thought and theory, although developed in different directions and with a differing focus of attention.

Ellis' (1989) earlier model of information seeking behavior divided the information seeking process into six activities: starting, chaining, browsing, differentiating, monitoring, and extracting. One of Ellis' monitoring methods lightly covered information sharing. Ellis identified information exchange through informal collegial communication as an important method of staying current:

> Many of those interviewed used informal contacts to help them keep up to date. Some relied very heavily on such informal contacts to keep them abreast of developments, and others stressed the importance of such contacts. … Social scientists immersed in an area and familiar with others working in the area often rely on such contacts to bring news and information to their notice and in this way keep each other up to date. (Ellis, 1989, p.195)

The described mutual information sharing, not studied in detail at that time, is thus a part of the overall information seeking activity, which is reminiscent of the notion of "invisible colleges" (Kealey & Ricketts, 2014; Cronin, 1982; Price, 1963, and many others).

In the Computer Supported Cooperative Work (CSCW) area, the focus in information and knowledge sharing has tended toward collaboration: emphasizing tool building, enablement and use (collaboration technology and interfaces), and understanding the user experience of collaboration, tools and work integration.

(Scott, Graham, Wallace, Hancock, & Nacenta, 2015; Bjørn et al., 2014; Modi, Abbott & Counsell, 2013; Ackerman, Dachtera, Pipek & Wulf, 2013; Cheng, de Souza, Hupfer, Patterson, & Ross, 2003; Benford, Greenhalgh, Rodden, & Pycock, 2001; Bentley, Horstmann, Sikkel, & Trevor, 1995).

In the Knowledge Management area, the focus related to information sharing has been on (1) organizational and project processes (and systems) to encourage and facilitate the capture of both tacit and explicit knowledge (and information) by employees, and (2) the building of organizational knowledge/information repositories supported by the work processes (Witherspoon, Bergner, Cockrell & Stone, 2013; Yahia, Bellamine & Ghézala, 2012; Wang & Noe, 2010; Davenport, 2002). Communities of practice has also been an important concept across the organizational and management literature, and in knowledge management (Pattinson & Preece, 2014; Wang, Yang & Chou, 2008; Davenport & Hall, 2005).

2.2.4 Information sharing: a research domain of its own

Information sharing has emerged as a primary research topic (Pilerot, 2014; Widén & Hansen, 2012; Pilerot & Limberg, 2011; Wilson, 2010; Talja & Hansen, 2006; Rioux, 2004) with a focus on various dimensions, methods, theoretical frameworks, and study settings. In this emergent time, even some articles with information sharing in the title (e.g., Talja & Hansen, 2006, and others) have focused more on collaborative information behavior rather than specifically on information sharing.

Rioux (2004) positioned information sharing on an equal status with information seeking in his Information-and-Acquiring theory. With a focus on web-based information acquisition and sharing, Rioux's work related to information encountering (Erdelez, 1997), as people often share information they unexpectedly find. He identifies additional related research areas, specifically:

> Future development of the IA&S concept may include: studying IA&S among specific groups (e.g., teachers, expectant mothers, scientists), exploring "non-sharing" behaviours, and examining this IA&S from the perspective of the receiver. An in-depth study of the relationships and communication channels between information sharers and receivers would also be valuable (Rioux, 2004, p.172).

Table 2-1 provides a tabular summary of meta-reviews and individual articles about empirical studies about information (including ones where information sharing is not a primary focus but include insights about information sharing or knowledge sharing). Table 2-2 provides a summary of the theories and theoretical frameworks used in information sharing-related research. Specific findings from these studies about information sharing factors are discussed in later chapters, and mapped to activity system constructs (Table 3-2) and findings from this research (Table 5-1).

Talja and Hansen (2006) classified a set of nine empirical studies from 1993 to 2005 according to their defined collaborative information behavior dimensions (p.124). Their ten dimensions mapped five pairs of contrasting characteristics that emerged from their review of these studies: synchronicity (asynchronously-synchronously), location (co-located-distributed), coupling (loosely coupled-tightly coupled), group (intragroup-intergroup), and connection (direct-indirect). These characteristics are also important factors in information sharing and the associated studies are foundational literature elements for this study especially in the context of distance. This body of work influenced the research design for this study, and the development of the Information Sharing Distance Model and the Information Sharing Discrete Distance Metric

2.3 INFORMATION SHARING AND DIMENSIONS OF DISTANCE

The next sections detail previous research about information sharing and the six specific forms of distance relevant to this study. These provide a good foundation for renewed conceptualization of information sharing over distance with multiple co-occurring distance dimensions taken into consideration.

2.3.1 Information sharing across geography

Virtual teams and dispersed work activities across physical distance -- particularly software development and design teams -- have been studied to a great extent. Distributed cognition (Hutchins, 2000) provides a framework for workplace studies in a variety of settings and with multiple factors. While information sharing is not always the primary focus, the role of information sharing often emerges as an important factor in these studies.

Common ground is identified as a critical factor in collaboration and teamwork in a geographically dispersed configuration in several studies (e.g., Bjørn et al., 2014; Olson & Olson, 2000; Clark, 1996), along with the concepts of "coupling of work, collaboration readiness, collaboration technology readiness, and organizational management" (Bjørn et al., 2014, p.1). Cramton's (2001) foundational work pointed to five issues that hinder common ground, which suggest the pivotal role that information sharing may play with regard to collaboration in a geographically dispersed virtual team. These information-related issues were: "(1) failure to communicate and retain contextual information, (2) unevenly distributed information, (3) difficulty in communicating and understanding the salience of information, (4) differences in speed of access to information, and (5) difficulty with interpreting the meaning of silence" (Cramton, 2001, p. 346). Other authors covered similar ideas using four different terms: situation awareness (Seebach, Beck & Pahlke, 2011; Sonnenwald, Maglaughlin & Whitton, 2004), collaborative grounding (Hertzum, 2008), group awareness (Scott, Graham, Wallace, Hancock & Nacenta, 2015; Gutwin, Penner & Schneider, 2004), and mutual knowledge (Koppman & Gupta, 2014; Cramton, 2001).

The importance of information sharing was emphasized to build collaboration strength, shared understanding, and common ground, particularly when not co-located (Zahedi & Babar, 2014; Hinds & Weisband, 2003). Also highlighted was the particularly troublesome configuration of a partially dispersed workgroup, where some people are collocated together and others are dispersed a distance away from that group (Siebdrat, Hoegl & Ernst, 2014; Voida, Bos, Olson, Olson & Dunning, 2012; Ocker & Hiltz, 2012). The lack of a level playing field across the workgroup members causes some imbalances between members. It is especially problematic to organize/conduct meetings with some people gathering around a table and others connecting by phone and/or a web teleconferencing device. One solution noted is to use a variation of "hybrid" configuration and have periodic face-to-face meetings with a geographically dispersed workgroup (Fiol & O'Connor, 2005). Another option to reduce the issues of large geographical distances and time zone differences was to set up a proximate location for offshoring activities (Carmel & Agarwal, 2001; Carmel & Abbott, 2007), with recommendations of how to manage distance issues (such as reducing intensive collaboration, cultural distance, and temporal distance). The approach of reducing intensive collaboration is contradictory

Author	**Year**	**Study Population**	**Focus**	**Research Design**
Pilerot	2014	Nordic network of design researchers	Documents as multidimensional objects; Nordic network of design researchers	Ethnography, survey
Sharp, Giuffrida and Melnick	2012	Small software development team	Information flow in dispersed agile team	Ethnographic case study; distributed cognition
Hassan Ibrahim and Allen	2012	Oil rig emergency response team	trust and information sharing in critical incidents	Activity theory; review of documents, observation, semi-structured interviews
Grubb and Begel	2012	989 Microsoft engineers	Dependency, sharing	Survey
Mishra, Allen and Pearman	2011	UK emergency responders	observation of 35 hours of multi-agency training and exercises; 20 semi-structured interviews (p.2)	Activity theory
Mesmer-Magnus et al.	2011	94 published studies; 79 student studies	Meta review	Review
Haussler	2011	1694 bio-scientists in 2 environments: academia and industry	Comparison of sharing likelihood; open science, reciprocity	Survey, social capital theory
Wang and Noe	2010	Review of 79 studies	organizational context, interpersonal/team characteristics, cultural characteristics, individual characteristics, motivational factors (p.117)	Review
Robinson	2010	78 engineers carried PDAs for 20 working days; logged activity every hour	Percentage of time performing different tasks	Quantitative analysis
Goh and Hooper	2009		Military, trust	Survey, quantitative analysis
Haeussler, Jiang, Thursby and Thursby	2014		Game-theoretic models of sharing	Model and survey
Mesmer-Magnus and DeChurch	2009	Review of 94 published studies		Meta-review
Dearman, Kellar and Truong	2008	20 paid participants in Toronto, Canada 18 years – 55 years of age	Using weak ties for collaborative information sharing in everyday life	diary study of everyday information needs/sharing
Suthers, Medina, Vatrapu and Dwyer	2007	students	Computer Science/ CSCL- utilized Stasser hidden profile	Experiment – hidden profile
Widén-Wulff and Davenport	2007	Finnish insurance and biotech companies	Link between information and org knowledge production	Case study
Talja and Hansen	2006	Review article – studies from 1993 -2005	Social Practice of CIB Studies	Review
Razavi and Iverson	2006	Interview of 9 high school students	Factors of privacy and trust in a personal learning space. Factors included sharing preferences, type of information, and purpose.	Grounded theory
Sonnenwald	2006	Participants in 3 military exercises (battlefield simulation, post-simulation review sessions)	Situation awareness, barriers to information sharing	Ethnography data (notes); formal and impromptu interviews (open-ended and critical incident questions)
Rafaeli and Raban	2005	Review article	Studies, theory, method, factors, variables for study: behavioral, social, and legal factors, and technological influences.	Review
Moye and Langfred	2004	135 MBA students in groups of 4 over 4 months working on multiple academic	Relationship between conflict, information sharing and group performance	Experiment

		projects		
Hirsh and Dinkelacker	2004	60 HP Labs Researchers	Information use to create new information	Survey
Ren, Y., Sha, C., Qian, W., Zhou, A., Ooi, B. C., and Tan, K. L. (2003)	2003		Experimental simulation of a peer-to-peer information sharing system using new algorithms inspired by small world phenomena	Experimental simulation
Miranda and Saunders	2003	32 groups of 5-6 undergraduates	Experiment - 32 five and six person groups with a fuzzy task	Experiment
Rafaeli and Ravid	2003	Student experimental study	Added email capability to computerized version of Sterman's production-distribution simulation, a standard supply chain operation simulation	Experiment
Miranda and Saunders	2003	Undergraduate information system students	Information sharing; decision marking; time constraints	Experiment
Churchill and Nelson	2003	Digital posterboards	Online and offline info sharing.	Fieldwork studies; quantitative and qualitative evaluations
Franz and Larson	2002		Impact of having an expert in a decision making meeting	
Talja	2002	Academic faculty; exploratory qualitative case studies	Information sharing related to document sharing in academic community	informal semi-structured interviews,
Olson and Olson	2000	Review article	Common ground, coupling of work, collaboration readiness, collaboration technology readiness. + organizational management (p.139)	Review
Millen and Dray	2000	Quantitative analysis of 4300 messages over 34 months; content analysis of 1800 msgs.- journalists LISTSERV	Creating and sharing collective goods; commitment and generalized reciprocity	Quantitative
Jarvenpaa and Staples	2000	Quantitative analysis of 4253 surveys by university employees	Perception of ownership	Quantitative analysis; uses Constant Theory of Information Sharing
Staples and Jarvenpaa	2000	Survey of university staff in UK and Australia	Perceptions that underlie use of electronic media	Quantitative
Graetz, Boyle, Kimble, Thompson and Garloch	1998	3 4-person groups; hidden profile.	Unshared information; effect of technology mediation; experiment – face-to-face, videoconference, instant messaging	Experiment
Constant, Kiesler and Sproull	1994	485 undergrad business majors	Vignette-based attitude measures of information sharing	3 experiments
Kleiner and Bouillon	1991	Quantitative analysis of 106 surveys plus public financial data	Quantitative correlation between firm data sharing policies and profitability	Quantitative

Table 2-1: Empirical studies about information sharing

Author	Theoretical components used in information sharing studies
Blomberg (2008)	Information transparency
Chatman (1996)	Theory of information poverty
Constant, Sproull and Kiseler (1996)	Weak ties in social networks; social theories
Constant, Kiesler and Sproull (1994)	Theory of information sharing
Dervin (1983, 1992, 1999)	Sensemaking
Graetz et al. (1998)	Hidden profile (Stasser)
Granovetter (1973)	Strength of weak ties in social network
Haeussler (2011)	Social capital theory
Haeussler et al. (2014)	Game-theoretic models of sharing
Hassan Ibrahim and Allen (2012)	Activity theory
Kim, Manley and Yang (2006)	Ontology based design framework, Communities of Practice, organizational learning
Miranda and Saunders (2003)	Social construction of meaning, social construction, social presence, and task closure theories
Mishra et al. (2011)	Activity theory; critical incident technique method
Moye and Langfred (2004)	Developed model of Information Sharing, Task Interdependence, General Mental Ability (GMA), Relationship Conflict, Task Conflict, and Group Performance
Poltrock et al. (2003)	Work analysis framework
Razavi and Iverson (2006)	Grounded Theory in Study of Information Sharing
Rioux (2004)	Information Acquiring and Sharing (IA&S)
Savolanien (2009)	Small worlds (Chatman), information grounds
Sharp et al. (2012)	Distributed cognition
Sharp and Robinson (2008)	Distributed cognition
Sonnenwald and Pierce (2000)	Situation Awareness
Stasser and Titus (1985, 1987)	Information Sharing Theory, Hidden Profile, Information pooling
Stewart and Stasser (1995)	Information sampling model
Talja (2002)	Social Networks
Widén-Wulff and Hansen (2012)	Social capital
Widén-Wulff and Davenport (2007)	Activity theory

Table 2-2: Theoretical frameworks used in information sharing studies

to Bjørn et al.'s (2014) recommendation of increasing interdependencies and collaboration. Holmström et al. (2006) investigated the three kinds of distance in combination: geographical, temporal, and socio-cultural, and found that agile practices did reduce the distance issues. However, despite these insights and strategies, crossing the gaps created by geographical distance remains a problematic area in practice.

2.3.2 Information sharing across time zones

Temporal factors (e.g., time zone) in work activities were often discussed in conjunction with geographical ones (Carmel & Agarwal, 2001; Carmel & Abbott, 2007), and there is indeed a very practical correlation: the farther the distance, the larger the time difference. Tang, Zhao, Cao and Inkpen, (2011) highlighted the issues experienced with significant time zone differences in a team (e.g., eight hours or more). And they noted, as did Olson and Olson (2000) that technology is not likely to resolve these issues, although use of technologies such as email can help to bridge the gaps caused by time zone incompatibilities. More recent studies (Shen, Lyytinen & Yoo, 2014; Wagstrom & Datta, 2014) confirmed that the temporal challenges of working across time zones are still a struggle.

2.3.3 Information sharing across disciplines

Studied since the early 1960s is the flow and exchange of information in the collegial networks of "invisible colleges", largely within disciplinary boundaries (Braun, Hefke, Schmidt & Sevilmis, 2007; Zuccala, 2006; Ellis, 1989; Weedman, 1983; Cronin, 1982; Price, 1963). Since that time, there have been many studies of individual populations and their information behaviors, with research subjects typically selected from a homogenous social sector. Examples of this are engineers (King, Casto, & Jones, 1994), students, and academic faculties (Talja, 2002), with a correspondingly homogeneous disciplinary background. Robinson (2010), for example, conducted a study of sampled work activities, where78 design engineers logged and categorized their information behaviors every hour for 20 hours. The study found that "substantially more time was spent receiving information they had *not* requested than information they had" (p.640), and confirmed earlier insights about the use of people as information sources (Robinson, 2010, p.655). Ehrlich and Chang, (2006) and Stasser, Stewart, and Wittenbaum (1995) also confirmed the utilization of people as sources of information.

Talja's (2002) study of academic faculty members focused on the social aspect of information, since sharing occurs with one or more additional people. The study noted four types of sharing: "strategic sharing, paradigmatic sharing, directive sharing, and social sharing" (p.1). The first three types of information sharing are specific to the academic setting, but could be adopted for other contexts. Social sharing is also generally applicable to information activities in everyday life as well as in specific work environments. Also noted were different modes of information sharing across disciplines and related to documents in social networks of an academic community, and highlighted the following activities of interest:

1. sharing information about relevant (and non-relevant) documents
2. sharing relevant documents
3. sharing information about the contents of relevant documents

4. sharing information about novel and efficient ways of finding relevant documents or information sources. (Talja, 2002, p.2)

Finally, Tabak and Willson (2012) conducted a study of an academic community and found that information sharing practice and context jointly shape each other.

Studies of inter-disciplinary or intermingled multidisciplinary populations were few in number (Reijonen & Talja, 2006; Sonnenwald, 2006; Maglaughlin & Sonnenwald, 2005), although some comparative studies between homogenous populations exist in the literature. Ellis and Haugan (1997) studied the role of information and information seeking behaviors of engineers and research scientists with the intent of creating a model. They found some significant differences in information seeking behaviors between the engineers and scientists.

> The engineers made heavy use of internal communication within their own departments or project teams or within the company… (They) chose their information channels based on their own experience and knowledge, through the consultation of personal contacts, or both methods. They had little experience of the use of information services on the network. (p.401)

The scientists depended on external colleagues more than the engineers, and also relied on librarians, alerts, and computerized searching to stay current through the scientific literature. Hirsh and Dinkelacker (2004) studied corporate researchers and highlighted the finding that information seeking is critical in order to create information. They also found that information sharing was an important associated activity:

> When asked via open-ended questions about their information-seeking- and-producing activities… (they) pointed to the value of cross-lab interactions, and having tools that would enhance communication and information sharing across space and time. (p.816)

These studies suggest several interesting observations about information sharing. Although there are common foundations, information behaviors by people of different disciplines or job categories have different preference and some variation. Information sharing is often an outgrowth of information seeking and information creation (on a continuum of different information behaviors). It is clear that information sharing across disciplinary boundaries presents particular issues. Both areas need further research. These findings suggest that additional studies are needed looking at the information behaviors of mixed populations, particularly information sharing, in order to build an integrated understanding of the phenomenon as it occurs in a heterogeneous setting.

2.3.4 Information sharing across organizational boundaries

Studies on cross-organizational information sharing and knowledge sharing are also plentiful. These studies focused on a wide variety of dimensions including individual characteristics, leadership, team, organization, social networks, motivations to share, and rewards in sharing, to name just a few (Pentland, 2014; Wang & Noe, 2010; Widén-Wulff & Davenport, 2007; Bock & Kim, 2002). The concept of organizational gatekeepers managing information flow (Allen, 1984) was from a time of less participatory and transparent organizational work environments, but it may still have a conceptual place today.

Allen and Henn (2007) divided communication of engineers and scientists into three categories: (1) communication for coordination, (2) communication for inspiration, and (3) communication for information (p.26). The third category, “communication for information”, included information sharing activities in an organizational setting, and this thread was picked up by Widén-Wulff and Davenport (2007) in their study of organizational knowledge development and sharing in commercial work settings using activity theory

Sonnenwald (2006) focused on organizational barriers to effective information sharing in an important study of military battlefield simulation, finding that a common understanding often does not accompany shared information between people. In this “command and control” context, effective information sharing could be a life and death matter, yet she found that barriers to effective sharing exist. By focusing on instances of “joint communicative action” for incidents of misunderstanding, Sonnenwald revealed more about the phenomenon and associated information sharing challenges:

1. Recognizing differences in the underlying meanings of shared symbols. (p.6)
2. Sharing implications of information. (p.7)
3. Understanding the role of emotions in sharing information. (p.9)
4. Re-establishing trust. (Sonnenwald, 2006, (p. 10)

Sonnenwald (2006) noted that these challenges "are influenced by inter-organizational, inter-cultural, and inter-disciplinary differences which emerged in both face-to-face and remote communication" (p. 12). An earlier work by Sonnenwald (1995) on contested collaboration within a design team highlighted the important roles of information behavior, framed as "observed communication roles that span group boundaries" (p.867). This includes a gatekeeper who performs a filtering/blocking process to and from the group, and a boundary translator who "translates group information for others who are not members of the group" (p.867). Highlighted here are communication problems that hinder collaboration and shared understanding, which also reflect information sharing issues.

The knowledge management literature also considered the organizational setting through a process-focused lens, with examples of studies that considered the effects of heterogeneous organizational aspects (Akoumianakis, 2014a; 2014b), the sharing effect of the social environment and management support for information sharing (Connelly & Kelloway, 2003), and the information sharing implications of structural factors such as organizational flattening and decentralization (Barua & Ravindran, 1996).

2.3.5 Information sharing across heterogeneous project roles

The specific job roles of team members in a project offer another dimension that contributes a vantage point on information sharing. This category may seem to be an overlap with disciplinary background, but in practice, there may be asymmetries or nuances between the educational backgrounds and their professional roles of each team member. Again, studies of homogenous populations are plentiful: in general (Golovchinsky et al., 2009; Hansen & Järvelin, 2005; Bunderson, 2003; Bunderson & Sutcliffe, 2002; Fidel et al., 2000); of academics (Tenopir, King, Spencer & Wu, 2009); of engineers (Allard et al., 2009; Hansen & Järvelin, 2005; Fidel & Green, 2004; Fidel, Pejtersen, Cleal & Bruce, 2004; Tenopir & King, 2004; Cool & Xie, 2000; Bruce, Fidel, Pejtersen, Dumais, Grudin & Poltrock, 2003); and of software engineers (Ko, DeLine & Venolia, 2007; Milewski , 2007; Milewski, Tremaine, Egan, Zhang, Köbler & O'Sullivan, 2007; Hertzum, 2002; Hertzum & Pejtersen, 2000). "Communities of Practice", a predominantly industry-based approach of aligning "like" members of a work activity, with information sharing as a primary part of those initiatives (Wanberg, Javernick-Will, Chinowsky & Taylor, 2015; Pattinson & Preece, 2014; Davenport & Hall, 2005; Pan & Leidner, 2003) is also a focus for studies of information sharing.

2.3.6 Information sharing across project tenure variations

This dimension is about the cognitive distance/organizational memory between newer and existing workgroup members when someone joins a project effort already underway. Tuckman's organizational dynamics (Bonebright, 2010; Tuckman & Jensen, 1977) addressed the creative spike small teams can experience in a membership change. Summers, Humphrey and Ferris (2012) looked at the effect of membership change related to project coordination and transfer of specialized knowledge during the change. However, literature is scant that addresses the broad information dynamic of team changes in a project, particularly in the dynamic and loosely coupled organizational settings in evidence today.

2.4 GAPS IN THE INFORMATION SHARING LITERATURE

The previous section detailed the solid foundation of literature related to information sharing over the past approximately twenty years. Nevertheless, it is also clear that additional knowledge can improve information. We need to know much more about how information sharing occurs in a variety of complex situations: in current organizational settings and configurations; with current technologies; and with current collaboration practices and ways of working. Scholars across the disciplines have called for additional study of the many facets of information sharing to address the gaps in the literature. Wilson (2010) noted that information sharing is an underexplored area of information behavior and called for application of more comprehensive theoretical frameworks, such as activity theory. Wang and Noe (2010) called for additional research in online vs. face-to-face knowledge sharing. Widén-Wulff and Davenport (2007) advocated for studies of information behavior that incorporate organizational issues utilizing an activity theory framework.

There are a number of literature gaps in the information sharing area, which provides an opportunity for this research to make a scholarly contribution:

1. This study has a primary focus on information sharing as an information behavior. An increasing number of studies in the past few years have information sharing as the focus, but the number is still relatively small. Many more include information sharing as an ancillary subject or incidental phenomenon. While that does build a base for further investigation, it is important for studies to keep information sharing front and center. This focuses squarely on information sharing, while incorporating relevant literature on information sharing as secondary or incidental findings.
2. This study focuses on information sharing as an information behavior within the Library and Information Science literature. Due to the ubiquity and pervasiveness of information sharing (or knowledge sharing) in many disciplines, there is a richness of cross-disciplinary perspectives about information sharing. This provides breadth in the understanding of information from multiple vantage points, but does not contribute depth situated on an information behavior foundation, and with an information behavior perspective. Information sharing deserves an information behavior lineage, just as long and as deep within the subdiscipline as information seeking has. This study primarily builds on information behavior literature, while incorporating relevant literature from other disciplines.
3. This study is an empirical study of information sharing utilizing broad, significant theories. Wilson (2013, 2010) has taken a leadership position in challenging the information science community to apply theories such as activity theory to the study of both information behavior in general and information sharing in particular. Needed are more studies that are empirical, although some studies in recent years have answered this call. The use of activity theory in this study enables a focus on the broad landscape of a complicated work activity, where the global distribution of professionals, often across varying time zones, as well as the disciplinary, organization, and job role dimensions, is quite complex. Activity theory also enables the analysis at a desired level of granularity. It provides the

best of both worlds: a broad perspective and detailed analysis. This study answers Wilson's call to utilize larger, broader theories. As Wilson (2013) notes:

> …. Engeström's formulation of activity theory offers the information behaviour researcher a framework that will enable the development of a coherent statement of the nature of the problem to be investigated, and will allow the researcher to ensure that the full scope of relationships within the activity system is understood. (p.26)

4. This study is an empirical study of information sharing between/among non-student participants. While short-term studies of students are helpful, more studies of information sharing among professionals in "production" settings, such as business, medicine, and other complex environments, reveal details of situated human activity in a particular environment. Some phenomena emerge only in a work setting and in a period longer than a semester or quarter enables. This study shines a light on real information sharing in a complex endeavour.
5. This study is an investigation of information sharing with a timeline long enough to observe changes. It is a challenge to sustain the focus of study in a real-life setting of sufficient duration to gather data on natural changes, growth, etc. Collection of data about a 16-month timeline of project events and changes provides a rich opportunity to build an understanding of embedded information sharing in a work context over a longer period.
6. This study focuses on multiple dimensions of distance in the same situated environment.
7. This study focuses on two understudied dimensions of distance in information sharing: heterogeneous roles in a work activity, and the effect of varying project tenure.
8. This study is an investigation of information sharing with the application of the activity theory theoretical framework to this highly specific setting of research software development in an industrial setting.

In short, this study brings many contributions to the fields of information behavior and information sharing.

2.5 ACTIVITY THEORY: THE THEORETICAL FRAMEWORK

A wide variety of disciplines has utilized activity theory over the past 20 years to explain multiple phenomena. The fields of human computer interface research and computer-supported cooperative work were early adopters of activity theory (Nardi, 1996). There have been a number of recent calls in the information science literature to broaden the theoretical foundation of research on information behavior (Allen, Karanasios, & Slavova, 2011), to accommodate individuals along with their social contexts and the technology dimension, and to provide better links to practice as activity theory does. Wilson (2013, 2008) has been an early and influential voice advocating the use of activity theory in studies of information seeking. In his review of activity theory research, Wilson criticized the conduct of library and information science research in "silos" and use of narrow theories. He asserted that the lack of broad theories limits the impact of the research across the field, and he advocated utilization of broader theoretical constructs, such as activity theory, to overcome this disciplinary weakness in library and information science research. Increasing focus on the use of activity theory in Information Science research is demonstrated by treatment of this theory in the *Annual Review of Information Science and Technology* (Wilson, 2008) and a special issue of the journal *Information Research* (Wilson, 2007) with half of the articles focused on activity theory. The *Journal of the Association for Information Science and Technology (*formerly the *Journal of the American Society for Information Science and Technology)* has published a steady number of articles related to activity theory in recent years. A sampling of these articles include: Spasser (1999); Xu (2007); Xu and Liu (2007); Stvilia and Jörgensen (2010); Allen (2011); Allen et al. (2011); Huang, Stvilia, Jörgensen, and Bass (2012); Sun (2012); Mervyn and Allen (2012); Goggins, Mascaro, and Valetto (2013); Isah and Byström (2015); Mishra, Allen, and Pearman (2015); Stvilia, Hinnant, Wu, Worrall, Lee, Burnett, Burnett, Kazmer, and Marty (2015).

The theoretical framework of this study is activity theory, adopted in response to Wilson's call for empirical studies based on broad theories in the library and information field, and inspired by the previous studies noted above. This decision was made with the following considerations: (1) it is a good fit with the research problem (information sharing in a workgroup with multiple dimensions of distance and a secondary focus on collaboration), (2) to add empirical evidence about information behavior and collaboration in an industrial setting, which can be difficult to access, and (3) to study more broadly the manifestations of distance in information behavior in distributed, cross-disciplinary workgroups using a unit of analysis (activity) with a wide aperture. This study of information sharing -- an important aspect of information behavior that is more recently emergent than some others are (e.g., information seeking) -- bridges a gap in the existing literature. Through this more general theoretical lens, it broadens our understanding about the rich dimensionality of information sharing in a complex technical context.

2.5.1 What is activity theory?

Activity theory has evolved over the past century since Soviet psychologists developed it in the early 1900s, although not published in the West until the 1970s. Closely linked with Marxism, it provides a systemic visualization and abstraction of complex systems involving human activity, and provides a broad contextual model for understanding the interactions and interplay among people (subjects), the human objective of the activity (object), and the tools, mediating artifact) in a larger social, historical, or work context (Widén-Wulff & Davenport, 2007). It simultaneously enables analysis at the level of the individual and larger collections of people such as an organization or community.

> Activity theory is a psychological theory developed over the course of some 70 years in the Soviet Union. It is concerned with understanding the relation between consciousness and activity and has labored to provide a framework in which a meaningful unity between the two can be conceived. (Nardi, 1996, p.xi)
>
> Activity theory incorporates strong notions of intentionality, history, mediation, Collaboration, and development in constructing consciousness. (Nardi, 1996, p. 7)

Definitions for activity theory constructs are contained in the front matter definitions section.

2.5.2 Activity theory background and development

Activity theory has its roots in a Soviet branch of psychology called "Cultural-Historical Psychology", developed by Soviet psychologist Lev S. Vygotsky in the early twentieth century and influenced by Marxism (Vygotsky, 1978). Despite limited access to accurate translations of his work by Western scholars (van der Veer & Yasnitsky, 2011), his ideas about language acquisition in children and the mediating role of language, writing, mathematics, and other symbol structures (Wilson, 2008, p. 121) continue to be influential in the scholarly community of psychology (Toomela, 2000). An adaptation of the graphic representation of his ideas in Figure 2-2 shows the learning experience of the human person through both language and direct interaction.

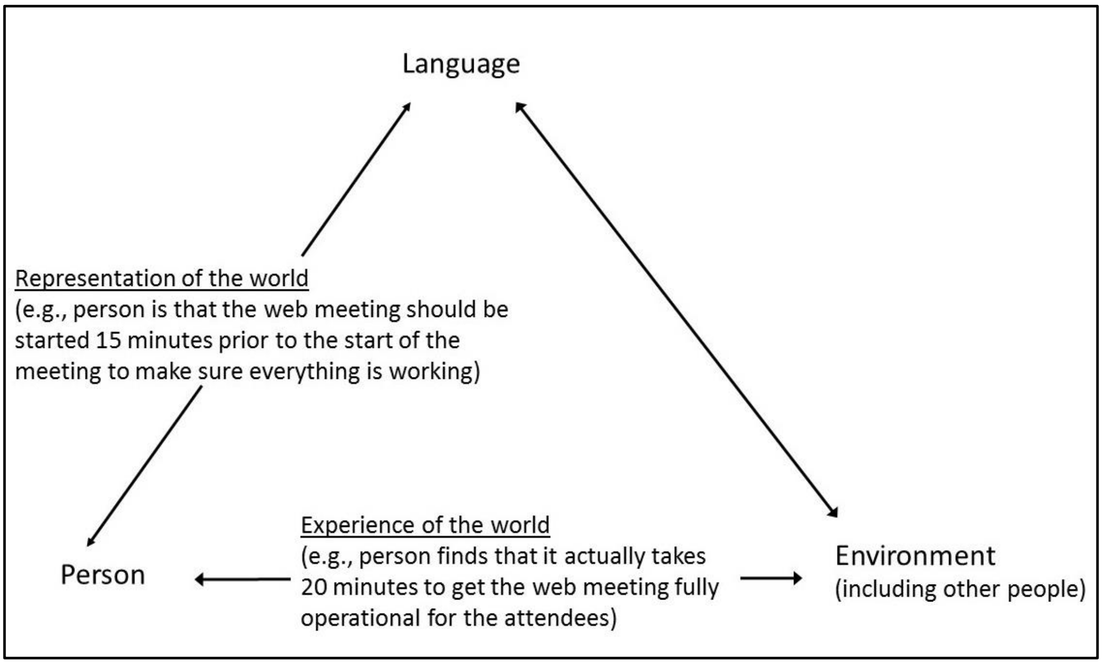


Figure 2-2: Mediation effect of language (Cole, 1993, p.19)

Considered the first generation of Cultural-Historical Activity Theory (CHAT), Vygotsky's work is the foundation for the development of on-going follow-on work: a second generation (Leont'v, 1974) and third generation activity theory (Engeström, 1987, 1999, 2001, 2007). That conceptual approach brought a systematic framework for studying the subject, the object, and the mediation of artifacts, the "mediating" effects of signs, language, and tools (Vygotsky, 1978; Cole, Göncü, & Vadeboncoeur, 2014). The concept of mediated action is the foundation for the first generation activity theory (Engeström, 2008a). Vygotsky's conception of mediating artifacts included abstract representations of mathematics, language, and symbols; and was expanded later by other researchers to include physical tools (Wilson, 2008). Illustrated in Figure 2-3 are the elements of Mediating Artifact, Subject, and Object. Important definitions of components in an activity theory context can be found in the front matter on page xvii, and include subject, object, and artifacts/tools, mediation, and knotworking.

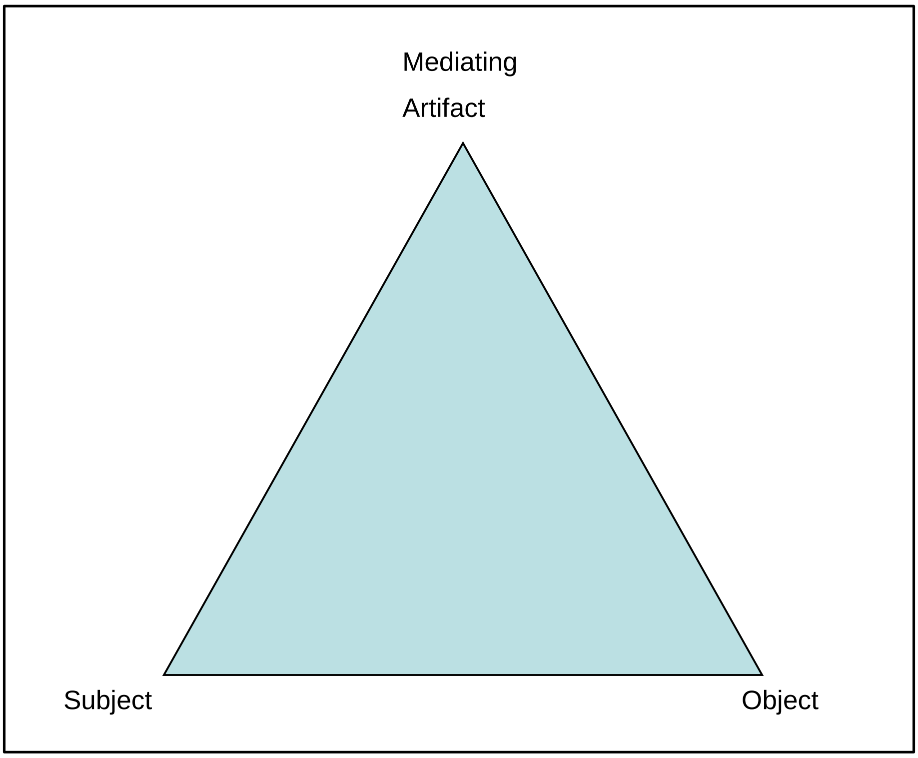


Figure 2-3: Vygotsky's activity theory model (Wilson, 2008, p.121)

Activity theory is more than a triangle, although the triangular diagram is a hallmark visual representation. The seemingly static diagram is really a snapshot, a model, a point in time, of an activity system that is in motion. Activity theory provides a strong context for analysis of complex, situated human activities, and has had a resurgence in recent years in the areas of Human Computer Interaction (HCI) (Kaptelinin & Nardi, 2006), Computer-supported cooperative work (CSCW) (Kaptelinin & Nardi, 2006), Software Development environments (Barthelmess & Anderson, 2002), and collaborative learning (Fjuk & Ludvigsen, 2001), as well as in Knowledge and Information Science research (Kuutti, 1996).

Continued theoretical work by Vygotsky's students (Leont'v, 1974) developed the cultural-historical dimension and the construct of division of labor (Wilson, 2008,

p. 121), and added the hierarchical concept of operations, actions, and activity (Wilson, 2013, 2008). Leont'v also emphasized the role of mediation in social relationships and introduced this influence into the activity theory model (Mills, Durepos, & Wiebe, 2009). The unit of analysis for this second-generation activity theory is the activity system (Engeström, 2008a). This extended the effect of mediation to include people as well as tools. Leont'v also provided definition for the unit of analysis at the level of the activity (Mills et al., 2009).

As Figure 2-4 illustrates, Engeström further extended the triangular model to include these new components: rules and norms, Vygotsky's community, and Vygotsky's division of labor (Wilson, 2008). These new components serve as additional mediating and interacting elements within the activity system, which also ground the model firmly in a larger community and social system. This expansion creates an even stronger model for studying activities in a work context, social issues in society, and a breadth of other topics and environments.

As humans, we accomplish tasks and objectives by engaging in an activity, by doing something tangible, something real. Nardi (1996) stated it very concisely:

> Activity theorists argue that consciousness is not a set of discrete disembodied cognitive acts…and certainly it is not the brain; rather, consciousness is located in everyday practice: you are what you do. And what you do is firmly and inextricably embedded in the social matrix of which every person is an organic part. (p.7)

Note the emphasis on *everyday practice* in the section above. Activity theory provides a broad frame within which to examine the intent of people (subjects) in real life settings as they work to accomplish something (object) using tools (mediating artifacts or instruments), in concert with other people (division of labor, community), by following (or not following) the rules and conventions in their natural environment (Barab, Evans & Baek, 2004).

Engeström (2001) has been instrumental in driving the continuing development of activity theory into its third generation, having introduced into activity systems the constructs of a network (Avis, 2009) and of partially shared objects (Engeström, 2008a).

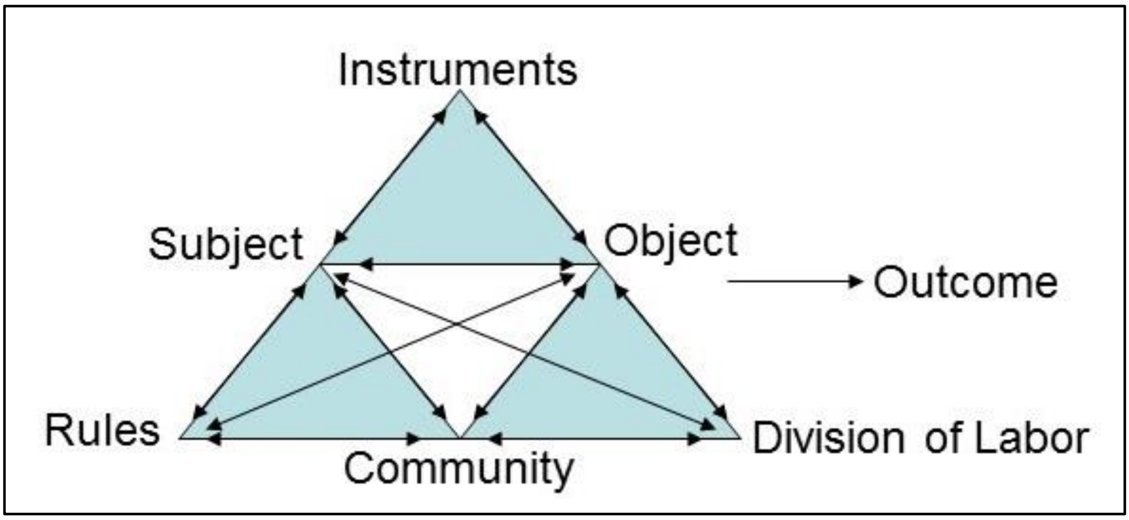


Figure 2-4: Engeström's second-generation activity theory model

Figure 2-5 shows an example of three interacting activity systems, each with independent subject (person or people), marked with an S in the figure, and with a shared object (marked with an O in the figure).

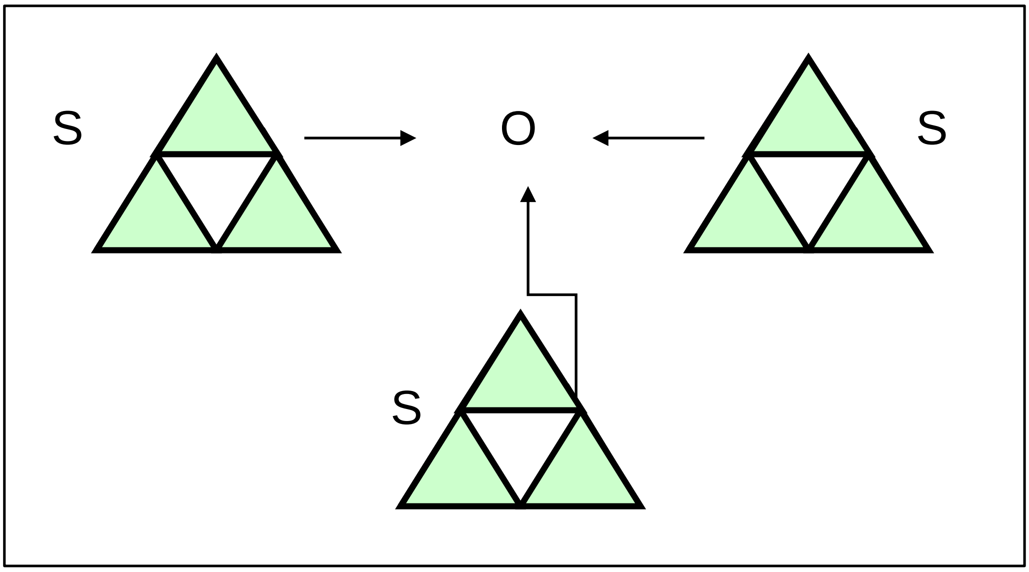


Figure 2-5: Engeström's third generation activity theory model

This third generation broadened the horizontal analysis capabilities, and contained the following five principles:

1. Unit of analysis: a collective, artifact-mediated and object-oriented activity system, seen in its network relations to other activity systems, is taken as the prime unit of analysis.
2. Multi-voiced: an activity system is always a community of multiple points of view, traditions, and interests.

3. Historicity: activity systems take shape and get transformed over lengthy periods of time.
4. Contradictions: play a central role as sources of change and development.
5. Expansive transformations: the movement of activity systems through relatively long cycles of qualitative transformations. This can also be thought of as the activity system itself moving through the zone of proximal development (see 3.3.2 in this document). (Engeström, 2001, pp. 136-137)

In 2006, Engeström further extended the model from the orderly networked linkage of multiple activity systems to a more dynamic one, using an organic metaphor of "mycorrhizae", an invisible living texture between a fungus and the roots of a plant (Engeström, 2007; 2006, p.1787).

2.5.3 Zone of Proximal Development

Vygotsky's concept of the Zone of Proximal Development is an important concept in activity theory, and the basis for further development of the theory.

> (The zone) is the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or collaboration with more capable peers. (Vygotsky, 1978, p. 86)

Wilson (2008) noted the importance of the Zone of Proximal Development concept, which is about how someone (a child or adult learner) can move to a higher level of understanding with the assistance of an adult or an advanced learner. Wilson observed that this benefit of collaboration or interaction with others may also be true in other contexts (p. 129), and empirical studies in diverse contexts have shown this to be true (Engeström & Sannino, 2010). Engeström (2000) also extended the zone into a space of transformation from an action to an activity.

In an activity system, an innovation may arise to address a contradiction or tension, as a manifestation of development or transformation in the activity system. This in essence becomes a Zone of Proximal Development, where multiple levels of contradictions ripple through an activity system, within and between elements, and drive development and change. Bødker (1991) further developed this idea and suggested that technical systems design can be informed and improved through activity system analysis and modeling to bring system creation and use back together.

Expansive learning is an activity theory construct that has the Zone of Proximal Development as one of its foundations, with the following properties:

> The theory of expansive learning puts the primacy on communities as learners, on transformation and creation of culture, on horizontal movement and hybridization, and on the formulation of theoretical concepts. (Engeström & Sannino, 2010, p.2)

In the dynamic context of this study, both the Zone of Proximal Development and Expansive Learning provide a lens to examine and understand the changes in an activity system over time.

2.5.4 Knotworking

One final dimension of activity theory with relevance to this study is a concept developed by Engeström and his collaborators, called knotworking. Knotworking is a dynamic style of connections between loosely coupled working groups, and unfolds in an activity system, or a network of activity. It draws on and extends the concept of co-configuration (Avis, 2009), and reflects the interdependent character that today's commercial enterprises have with customers, partners, and competitors (Engeström, 2008b).

> In knotworking, collaboration between the partners is of vital importance, yet it takes shape without rigid, prederminded [sic] rules or a fixed central authority. (Engeström, 2008b, p.20)

While it is true that the organizations in an enterprise do have a hierarchical reporting structure and a set of rules and commitments, knotworking provides a theoretical lens within the activity system theory to examine collaborations and interactions initiated by the members of a workgroup.

2.5.5 Criticisms of activity theory and directions for future development

No theory is perfect in representing the complexity of humans operating in a complex, real-life system, and activity theory is no exception. Lively debates in the literature have ensued about weaknesses in activity theory, and areas where scholars need to further develop the theory. Davydov (1999) identified eight problem areas of activity theory:

1. Understanding transformation
2. Collective and Individual Activity
3. Structure and components of activity (definition)
4. Different kinds of activity
5. Understanding communication
6. Connections to other theories
7. The biological and the social (relationship)
8. Organizing interdisciplinary. (pp. 42-49)

Related to some of these issues are philosophical questions about the nature of knowledge, to the political space, and to economics. The criticisms relevant to discuss here are the more pragmatic aspects in the application of activity theory in an empirical study such as this one. Awareness of these relevant concerns is a practical way to deal with them in an empirical study. Another area of criticism is weakness in representing power. Although the activity theory framework provides for hierarchical representations of organization, it has been criticized for not representing power adequately (Avis, 2009; Langemeyer & Roth, 2006).

Toomela (2000), focusing on the extension of activity theory from the original cultural-historical psychology foundation, expressed concern that "the analysis of an activity approach leads to serious doubts as to whether it is able to lead us to an understanding or explanation of mind or any specific psychological function" (p.353). He also asserted that activity theory, if separated from cultural-historical psychology, is a dead end. A subsequent article (Toomela, 2008) reiterated many of the same concerns and added methodological concerns in cultural psychology. Noted also as an issue was the failure to address issues of power and social antagonism (Avis, 2009). Engeström (2009) addressed the future of activity theory and these criticisms, with a call to continue the dialogue, to share data and studies, and to continue to address the issues and weaknesses of activity theory. He suggested five "mechanisms" for activity theory development to begin with in addressing these issues:

1. Living movement
2. Breaking away
3. Double stimulation
4. Stabilization
5. Boundary crossing (p. 312)

It will take time to realize these suggestions in further philosophical discussions, and for empirical studies to be completed. However, it is a healthy sign that activity theory is alive and relevant as long as these debates go on.

2.6 BACKGROUND TOPICS TREATMENT

2.6.1 Distance collaboration

Distance collaboration is a rich topical area in the literature (Pallot, Martínez-Carreras & Prinz, 2010; Fischer, 2005), with multiple disciplinary treatments, foci, and methodological approaches. For Computer Science, the ACM Computing Classification System places "collaboration in software development" (and a subcategory "programming teams") within the major topic area "software and its engineering".

The human activity of collaboration (Croker et al., 2009; Lauche, 2005) in a workgroup setting referred to the process of "thinking together" (Larsson, 2003, p.153), and in a virtual team setting about "thinking together apart" (Larsson, 2003, p. 153). The intellectual work and shared understanding of software design and development requires collaboration. One definition of collaboration identified six key activities of collaboration as communication, information sharing, coordination, cooperation, problem solving, and negotiation (Croker et al., 2009, p.32). All of these activities occur in software engineering projects, and their individual and cumulative effectiveness will either build toward overall success for the project, or detract from it. Serce, Alpaslan, Swigger, Brazile, Dafoulas and Lopez (2009) looked at the triggers for collaboration in software development teams and found that:

> …the communication categories with the highest number of activities were organizing work, initiating activities (two planning behaviors), feedback seeking, feedback giving (two contributing behaviors), and social interaction. There were very few communication activities tagged as (a) reflection and monitoring, (b) challenging others (a contributing subcategory), or (c) advocating effort (a seeking-input subcategory) (p.495).

Research on virtual teams emerged as soon as people started experimenting with collaboration over the Internet in early 1990s. Powell et al.'s (2004) critical review of early studies on virtual teams (published from 1991 to 2002) revealed that the vast majority was less than a semester in length (29 studies vs. 13 studies) and used students as research subjects. Slightly more than half of the studies were non-global, meaning virtual teams composed of people from one country or area. Important factors in the virtual team literature cited by Powell et al. (2004) include: inputs (design, culture, technical, and training), task processes (communication, coordination, task-technology-structure fit), socio-emotional processes (relationship-building, cohesion, and trust), and outputs (performance, satisfaction) (p. 8).

Powell et al. (2004) also identified a number of issues in association with the operation of virtual teams, particularly in contrast with a traditional proximally co-located configuration. These include difficulties with communication, trust, work processes, cohesion, conflict resolution, mutual knowledge, coordination, training, and comprehension, to highlight just a few. Powell noted the fact that traditional teams outperformed virtual teams in "orderly and efficient exchange of information" (p.8). And "face-to-face" meetings are advocated as one intervention to improve the operation of a virtual team.

2.6.2 Virtual teams in software engineering

Establishment of virtual teams, particularly ones that are globally distributed, has been particularly dominant in the software engineering arena. Software engineering teams that are "virtual" (i.e., geographically dispersed) experience all of the challenges and issues noted for virtual teams, some to a greater degree with addition of a few more. Lack of awareness within the group of what other group members are doing when there is no physical proximity is a major problem which ripples throughout the software engineering process (Herbsleb, 2007; Gutwin et al., 2004). This can be due to uneven communication and information dissemination procedures, or because of a deliberate decision to withhold information from others. Development of integrated development environments (Hupfer, Cheng, Ross & Patterson, 2004) -- which include collaboration capabilities in a consolidated computing environment – was noted as an attempt to utilize technology to address this shortcoming.

Another major difficulty highlighted of virtual teams was the issue of project management with a global software engineering team. This "coordination over distance" (Herbsleb, 2007; Carmel & Agarwal, 2001) is a stress to standard project management methodologies and practices. In addition, because individual programmers and architects are each building parts for the completed whole, working on a *common ground* is critical to having a high quality operational product at project completion.

Information and collaboration are two core characteristics of software engineering. Software engineering is an inherently information-intensive activity, both on an individual basis and from a collaborative perspective. Software developers are daily consumers of information produced by others, such as project status reports, project plans, and even software codes. They also produce information which is shared both formally and informally with their colleagues, e.g., interpretations of user requirements, and specifications for user interface.

In a large team configuration, software engineering is a particularly interdependent activity, since the software produced by each developer is integrated to work seamlessly with the code of others. Difficulties may arise if information is not shared, shared unevenly, or shared with a different time dimension, because of either lack of proximity, varying time zones, or differing computing environments (Cramton, 2001). Each workgroup member performs tasks which may impact others' tasks, change the technical environment, and convey information to others about their work. Each workgroup member also draws on personal expertise of intrinsic (and sometimes unique) domain knowledge. It is important for workgroup members to share task-related and contextual information. If there are gaps in, or barriers to, sending or receiving this information, particularly in the characteristically fast-paced environment today, it will have a negative impact on the project and on the collaboration. Finally, each workgroup member can serve as a producer, consumer, and disseminator of information, with varying scope and impact through both formal and informal roles (Prekop, 2002).

Many experts and researchers have explored technical process, environment, and management of software engineering over the past 25 plus years with the objective to improve project success rates. Their suggested approaches for improvement included better software estimation techniques, integrated software engineering environments, and improvements in requirements and specifications (Verner, Overmyer & McCain, 1999; Hupfer et al., 2004;

Belkhatir & Ahmed-Nacer, 1995). However, the role of information sharing and use by heterogeneous and multi-disciplinary project workgroups has not been adequately studied. Members of the project workgroup serve in adjunct roles (such as project management, quality assurance/testing, customer representatives, and line management), and at the same time can be dispersed geographically. Improvement of their practices of information sharing and information use ought to be included as part of the solution as well.

2.6.3 Agile software engineering methods

Finally, just a few words about the agile family of processes in software engineering, and what it brings to this study of information sharing. Agile methods provide an alternative to the documentation and process-focused methods of earlier approaches, with the integrated (and operational) software code as the object of focus (Zaitsev, Gal & Tan, 2014). The code is the thing.

Scholars are divided about the impact that agile methods have on distance collaboration, with some finding that agile caused problems (Shrivastava & Date, 2010), while others argued that agile processes solve issues in virtual configurations (Beecham, et al., 2014; Holmström et al., 2006). Many scholars advocated additional studies with a granular focus to sort out these differing viewpoints. This study, while focusing primarily on information sharing, does provide observations about an agile method in an environment of multiple distances.

2.7 SUMMARY

This chapter reviews the overall landscape of the study, along with a detailed view of important features. The treatment of information sharing (and knowledge sharing) in the literature across multiple disciplines is analyzed, with a focus on information behavior within the library and information science field, as well as in the areas of computer supported cooperative work (CSCW in Computer Science), and business, management, and organization. Next is a detailed discussion of information sharing in co-existence with key dimensions of distance in collaboration: geography, time zones, organizational boundaries, heterogeneous project roles, and project tenure variations. Gaps identified include the need for additional empirical studies of information sharing with information sharing as an information behavior being the primary focus, studies with sufficient longitudinal length, studies that utilize larger and broader theoretical frameworks, studies in real-life, complex environments, and studies on information sharing that explore the impact of work efforts of people performing heterogeneous roles, and of team members joining a project after it has started. Discussed are the foundations of activity theory, with arguments for the strength of activity theory to illuminate insights and nuances of information sharing. Last discussed are some background components to put this study in context: distance collaboration, virtual teams in software engineering, and agile software engineering methods.

Chapter 3: Research Design

3.1 CHAPTER OVERVIEW

The previous chapters present a high-level overview of the phenomena and challenges of teamwork over distance, and a critical review of the scholarly literature in the areas of virtual teams, information sharing, and collaboration in various contexts, particularly in software development projects. The overall research aim for this study is to gain insight on what people do when they share information, how they think about sharing, what they share, what others share with them, and what this all means for collaboration. This study examines information sharing activities of an extended project workgroup and stakeholders in a commercial enterprise. The purpose is to understand how those activities affect collaboration, along the continuum from building connections between workgroup members across geographic distance, incompatible time zones, heterogeneous roles, and dissimilar disciplines (on one end of the continuum), to exacerbating the separations that exist and creating further intra-workgroup challenges (on the other end). There is evidence in the literature that information sharing can be a positive factor in the (simpler) *coordination* of activities related to the project work effort (Bayerl & Lauche, 2008). This study extends this line of inquiry to examine the range of influences that information sharing has on the higher order *collaborative activity* in the project. In particular, building on Sonnenwald, Söderholm, Welch, Cairns, Manning, and Fuchs' (2014) and Maglaughlin and Sonnenwald's (2005) research on collaboration and information behavior, this study examines how information sharing affects factors such as common ground (an important factor in collaboration) among participants in a dispersed, multi-role, educationally heterogeneous, and highly technical workgroup.

This chapter describes the overall research design, with a focus on the theoretical lens and associated methods. First, the overall research philosophy and approach are presented, followed by a discussion of the core research question, the general problem space for the study, and the subordinate research questions to be explored.

Next covered is the application of activity theory, followed by the suitability of activity theory to answer the research questions of this study. An exposition of criticisms about activity theory provides a comprehensive perspective on activity theory and on the epistemological discussions about it in the literature. A review of other theories considered for potential application to this study, but rejected, provides an exposition of the rationale of choosing a theoretical framework for this study.

In addition, this chapter presents the approach of data gathering and analysis, built on the repertory grid technique (Kelly, 1955; Adams-Webber, 2006), and followed by a detailed description of the implementation of associated methods. This part also reports the two pilot studies done in advance, along with changes of the method of the main study because of learned lessons. Next, the methods of the main study are detailed. The detail of the ethics clearance information precedes a chapter summary.

3.2 RESEARCH PHILOSOPHY AND APPROACH

The foundational philosophy for this research is constructivist and interpretive, focused on the manifestations of the sharing of information in artifacts and events. As a reminder, Section 2.2.2 contains the working definitions of information and information sharing in this study.

Activity theory provides the grounding for this research in the tradition of Vygotsky, Leont'v, and Engeström. Activity theory provides a framework to understand the big picture as well as the small details. The activity system constructs enable the large view, and the capability to examine a large, complex, living, and constantly changing phenomenon: humans, individually and in concert with one another, expressing and enacting their internal ideas and motivations in real settings, utilizing language, technology, and other tools. It is a strength of activity theory to look at granular details while maintaining a high-level perspective. This results in a view of both the big picture and the small details.

The Zone of Proximal Development (Vygotsky, 1978) provides an opportunity for learning and subsequent changes in the activity system. Engeström and Sannino (2010) refer to this as expansive learning. Knotworking (Engeström et al., 2012) is a further development of activity theory and provides a framework for networks of people and workgroups collaborating in a loosely coupled configuration.

The focus is on building an understanding of information sharing and collaboration in the specific context of work activities as experienced and articulated by the participants. Such a specific context reflects their understanding and intent, as well as their organizational vantage point, their views as shaped by their educational and disciplinary backgrounds, and by their project role. Activity theory "is not a predictive theory" (Wilson, 2006, Conclusion), but does provide a general conceptual framework for understanding human activity. It is important to note that in this work activity context, the individual people, workgroup(s), and organizations all construct meaning and make sense of the environment, both individually and with other people. Their perspectives are contextual as individual and/or group experience, not an absolute truth. Therefore, the design of the research captures the perspectives of individual project workgroup members about information sharing and collaboration in their own words and vocabularies, through elicitation using the repertory grid technique and through examination of project artifacts.

Use of the repertory grid elicitation technique, a specific semi-structured interviewing technique, enables the participants to explain their view of the work context and activities using their own vocabulary. They name the mechanisms of information sharing (nouns), and describe the characteristics (adjectives) of those mechanisms. Section 3.4.2 provides a more detailed description of the repertory grid technique. Both activity theory and the repertory grid preserve the critical viewpoint and voice of the individual.

In the next phase, data and content analysis focuses deeply on their articulated accounts of information sharing and collaboration in the context of the work activity, to identify both common and unique patterns, and to perform a systemic mapping of key project events and activities. The project artifacts augment the interviews and undergo similar content analysis. Activity theory prescribes the unit of analysis at the level of an activity, meaning the human (subject) effort (operation, action) focused on achievement of an object in a specific context.

The goal of such analyses is to elucidate how the subjects see their work, their information, and project collaboration, along with information sharing activities in both directions (giving and receiving), what their colleagues are doing related to information sharing, and the accompanying motivations (the "why" associated with actions and activities). The participants may have multiple interpretations of information sharing events, artifacts, objectives, collaboration, and outcomes, along with multiple views on the role that information sharing plays in misunderstanding and success, and in building or damaging collaboration. This can lead to multiple views within a workgroup about resulting tension and contradictions, and opportunities for innovation.

In short, activity theory (via the associated analysis of activity systems) provides a comprehensive framework for gathering and analysing data from an interrelated set of participants (participants) about their work purposes and activities, helping to gain a view of common, collective objectives as well as individual, unique objectives. It provides the capability to situate each individual within the larger, complex "system" in which they work, as well as varying levels of "zooming out" to a larger landscape. In addition, activity theory provides a structure to understand changes that occur in that setting, including developmental changes, innovation, and contradictions that can drive innovation and change. Activity theory also provides helpful theoretical and conceptual models to characterize and understand the external manifestation at multiple levels of what people do through the constructs of operations, actions, and activity, from the most automatic requiring little or no thought, to increasing levels of inner cognitive processes expressed externally.

3.3 RESEARCH PROBLEM SPACE AND RESEARCH QUESTIONS

The core question of this research is: "How does information sharing occur in the distance collaboration of virtual teams?" The research problem centers on the phenomenon of information sharing in workgroups, and the associated challenges to collaboration and the accomplishment of work objectives by an extended project team. This research looks at how information sharing activities affect collaboration (e.g., bringing people together, acting as a source of misunderstanding) in the specific ways as indicated by the research questions.

Primary Research Question and Problem space: "How does information sharing occur in the distance collaboration of virtual teams?" The research problem centers on the phenomenon of information sharing in workgroups, and the associated challenges to collaboration and the accomplishment of work objectives by an extended project team.	
Subordinate Research Questions:	Research Questions of Talja and Hansen (2006, p.116):

1. How do information sharing activities manifest themselves in distance collaboration? 2. When and in what kinds of circumstances does information sharing occur in distance collaboration?	How do collaborative activities manifest themselves in IS&R? When and in what kinds of circumstances does collaboration in IS&R occur?
3. What types of information sharing behaviors and forms of shared information can be identified?	What types and forms of collaborations can be observed and identified? (e.g., collaborative browsing, searching, filtering)?
4. What attributes are related to different types and forms of information sharing in distance collaboration?	What attributes are related to different types and forms of collaborative IS&R?
5. What purposes does information sharing serve in distance collaboration?	What purposes does collaboration in IS&R serve?
	How should collaborative information sharing be accounted for in IR systems design

Table 3-1: Research problem space and research questions

The research questions in Table 3-1 are adapted from, or modeled after, the research questions of Talja and Hansen's (2006) study of Collaborative Information Behavior (CIB) and Collaborative Information Seeking and Retrieval (CIS&R). Listed in the right column are Talja and Hansen's (2006) original research questions, aligned in rows for line-by-line comparison. Talja and Hansen's (2006) research questions provide a proven framework of inquiry in the area of information behavior and collaboration, although their study had a different focus and objectives.

3.4 OVERALL RESEARCH DESIGN

Next is an exposition of the overall research design created to gather and analyse appropriate empirical data with the purpose of answering the identified research questions. Also detailed is an argument that activity theory in this case is the best approach to answer the research questions. The alignment between the research questions and the research design is important, as well as an understanding of the strengths and any weaknesses in the approach adopted.

3.4.1 Data gathering: Activity theory and the Repertory Grid Technique

Activity theory, and the repertory grid technique, has been widely used over many years across multiple disciplines, including the social sciences, human computer interaction, medicine, and sports (Jones, Edwards & Filho, 2014; Saúl, López-González, Moreno-Pulido, Compañ, Corbella & Feixas, 2012; Nardi, 1998; Kuutti, 1996). Activity theory provides an overall framework for this study with these characteristics: (1) a structure for in-depth look at one scenario, environment, or situation, (2) a framework for examining participant views from differing perspectives, (3) the capability to perform end-to-end analysis of activities and events, and (4) rigor in the case definition and data analysis.

The particular case (a virtual team of software development) chosen for this study is of interest for its manifestation of a *phenomenon* (information sharing in distance collaboration). The study is exploratory and descriptive, and the case serves the research objective well, which is to identify and describe information sharing as an important collaborative activity among the human participants.

Activity theory is the most suitable for this research for three reasons. First, Cultural-Historical Activity Theory (CHAT) provides a context for in-depth examination of many different aspects of human information sharing practices. CHAT enables a comprehensive and exhaustive examination of a single instance and can provide a view of subtleties in the situation. Secondly, the systematic analysis methods of CHAT provide an opportunity to examine information sharing practices from a comprehensive perspective, by analysing impacts, causes, and effects from the viewpoints of both the senders and recipients of information in the same context. This provides a holistic and comprehensive examination of interrelated components. Finally, the data is gathered and analysed in the complex context of a technical and organizational setting. Activity theory has a degree of integrity in observing the information practices of the project workgroup and gathering data in the "natural habitat" (work environment) of the project workgroup. The environment of the selected case is a real, industrial environment, and the data gathering/analysing methods provide a lens to examine this activity situated in the real world.

Repertory Grid Elicitation Technique

The repertory grid elicitation technique provides a specific structured method for the semi-structured interviews. Based on Kelly's (1955) work in Personal Construct Theory, the repertory grid technique has the perspective that each person is an individual scientist in his/her own world, and constructs a world-view and perspective that is unique. The repertory grid technique enables an open and systematic revelation of those internally held beliefs and perspectives. This technique has sound theoretical foundations, but is not a major theoretical framework in this study. The repertory grid elicitation technique is congruent with activity theory in carrying forward the subject-centered viewpoint. Repertory grid provides the capability for the participant to reveal their way of thinking, their vocabulary, and what different aspects of the discussed content mean to them with minimal or no guidance from the interviewer. The repertory grid technique, in contrast to a more directed style of questioning, brings the user terminology and perspectives cleanly into the data, where activity theory analysis advanced a similar user perspective.

The repertory grid elicitation technique was used in the semi-structured interviews to elicit responses from the participants, so that observations in their own words and worldview can be recorded about what is shared in the project (either what they share or what is shared with them) and how they experience those information sharing mechanisms in the project.

3.4.2 Data analysis: Activity system and content analysis

Analytical methods enabled making sense of the data in aggregate after the data gathering was completed. As activity theory is the theoretical lens to examine the phenomenon of information sharing, activity system analysis and modeling is the primary mechanism to analyse the data. This analysis technique provides a systematic approach to situate the components identified by each participant in a framework that can be compared and contrasted, while zooming in or out to greater or lesser levels of granularity. The hierarchy of mediating artifacts (Collins et al., 2002) provides a framework for the classification of identified artifacts.

In activity theory, an important foundational idea is that the unit of analysis is the activity system, as noted by Boer, van Baalen, and Kumar (2002):

> The activity theory emphasizes the importance of a systemic analysis of an organizational setting by considering it as (a network of) activities. With the activity system as a unit of analysis, the activity theory avoids simple causal explanations of knowledge sharing by describing an organizational setting as an ensemble of multiple systematically interacting elements (e.g., social rules, mediating artifacts and division of labor)….by taking the perspectives of different actors of an

> activity system, a system view is complemented with a subject's view. (p.1484)

The capability to preserve the voice, perspective, and explicit viewpoint of the participant while also creating activity systems from other vantage points is a powerful analytical approach to examining a complex system. The assurance of coherence and alignment from the earliest point of data gathering (i.e., the repertory grid elicitation interview) was important as the participants named the shared in their activity system: artifacts, mechanisms, and events. Activity theory paired with repertory grid elicitation is methodologically consistent from data gathering through data analysis. From the beginning of the study, the data collected captures the participant's unique viewpoint of how things look from his/her angle. Then in data analysis, activity theory provides the lens to look at the activity system of the individually situated person with the most detailed granularity, and at the same time offering views of less granularity (to varying extents) that provide more breadth.

Leximancer 4, a machine learning-based text analytics tool, provided automated thematic and concept analysis in conjunction with the activity system analysis and modeling. This provides broader insight across the interview transcripts, the individually elicited repertory grid elements (what is shared), and the constructs (adjectives describing what is shared), to augment manual reading and analysis through text analytics derived from the data and a visualization capability.

3.4.3 Suitability of activity theory to answer the research questions of this study

It is clear that activity theory provides a strong framework to examine and explain both the large landscape of a complex human situation and the small, granular details that comprise it. It is a great strength for a theoretical framework to accommodate both the macro and micro perspectives at the same time. In addition, Widén-Wulff and Davenport (2007) note that activity theory analysis enables analysis of a system in motion:

> By emphasizing mobility, fluidity, development, and learning, activity theory overcomes the limitations of more static task analysis, and provides a vocabulary to describe evolving knowledge production in terms of specific information behaviour. (p.8)

Allen et al. (2011) also argue for the advantages of activity theory in information behavior research:

> … In particular, CHAT provides researchers a theoretical lens to account for context and activity mediation and, by doing so, can increase the significance of information behaviour research to practice. (p.776)

Three areas in information behavior would benefit from the theoretical strength of activity theory, again according to Allen et al. (2011):

1. Balancing the "societist" and individual contexts,
2. Addressing the role of technology, and
3. Models to reconnect information behavior research with practice and policy. (p.777)

Widén-Wulff and Davenport (2007) note two major benefits in the use of activity theory in information behavior research: clarification of the term "information sharing" in specific contexts, and adoption of an analytical lens that encompasses the organizational context:

> Activity theory has expanded our understanding of information behaviour in two major ways. First, it has forced us to clarify our terminology. Many human information behaviour studies are confounded by indeterminate terminology; a point made recently by Bartlett and Toms (2005). By tying a term like 'information sharing' to a range of activities and actions whose salience varies across a number of organizational processes we are forced into specific usage… Secondly, as noted by Wilson, activity theory forces us to expand the horizons within which we observe and explore behaviour. Actions and operations are traced across different organizational processes, and sequences of inputs and outputs are made visible in ways that cannot be understood when research is based on more limited accounts of task-based work. (Conclusions section, para. 2)

In short, there is substantial information science literature providing a foundation for this choice to use activity theory in this study. The two main points are: (1) the methodological calls for broadening the theories used in information behavior (Wilson, 2013, 2008); and (2) the pervasiveness of activity theory studies across numerous disciplines provide evidence that this theoretical framework is sound for empirical study of a complex phenomenon (Engeström, 2008; Nardi 1998).

An examination of key information sharing factors gathered from the literature mapped against activity theory components (illustrated in Figure 3-1) shows the depth of the strength and suitability of activity theory for this study.

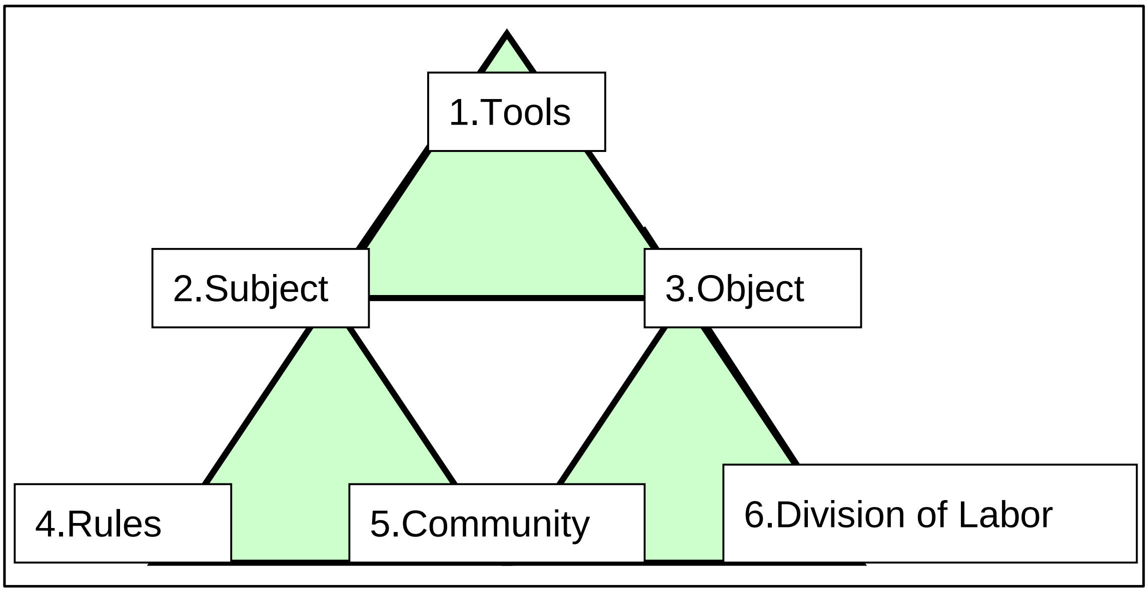


Figure 3-1: Activity system framework/components

Table 3-2 shows the mapping of factors related to information sharing identified from the literature to the activity system framework components, with the first column listing citation references, the second column listing factors identified in the corresponding studies, and the third column listing numbers which identify the matching activity system framework components, as defined in Figure 3-1. The factors/features in boldface type and the associated activity system component number(s) are a selective subset of identified factors from the literature that are relevant to this study.

This exercise highlights the strength of activity theory in providing a scaffolding to examine these diverse aspects of a complex human information behavior at a fine-grained level of detail, while maintaining the unit of analysis at the level of the activity. Other factors are included in the table for completeness, but are shown in regular type and do not have an activity system component number. Also shown is a list of types of information sharing in an academic context (Talja, 2002) with a similar mapping to the appropriate dimension of the activity theory system.

Reference	Factor/Feature of information sharing	Activity theory component
Pilerot (2014)	Documents as multidimensional objects	
Bjørn et al. (2014) (focus of this article is on collaboration, thus on validating these factors for information sharing)	**Common ground**	**2, 5, 6**
	Coupling of work (tightly, loosely)	**6**
	Collaboration readiness	
	Collaboration technology readiness	
	Organizational management (p.1)	
Sharp et al. (2012)	Complex digital artefacts	
	Information sharing has to be explicit (geographical distance)	**1, 2, 5, 6**
	Individual decisions regarding what/when to share	
	Technology difficulties	**1**
Hassan Ibrahim and Allen (2012)	Information sharing fosters trust	
Grubb and Begel (2012)	Organizational characteristics	
	Dependency perception	
	Information sharing attitudes	
Mesmer-Magnus et al. (2011)	Willingness to share unique information	
	Information sharing openness	
Wang and Noe (2010)	Organizational context	
	Interpersonal characteristics	
	Team characteristics	
	Motivational factors	
Robinson (2010)	**Active and passive information sharing**	**1, 2, 5, 6**
	Human sources of information	**1, 2, 5, 6**
	Non-human sources of information	**1, 2**
	Problem-solving and decision-making intertwined with information searching	**1, 2, 3, 6**
Wilson (2010)	Trust	
	Risk	
	Reward (or benefit)	
	Organizational proximity	**1, 2, 5, 6**
Goh and Hooper (2009)	Barriers to information sharing in a closed information environment; security	
Haeussler et al. (2014)	Specific and general information sharing in bio-science; competition.	
Suthers et al. (2007)	Integration of multiple info. sources and convergence on common solutions	
Widén-Wulff and Davenport (2007)	Individual and collective information behavior/decisions intersection with organizational processes.	
Sonnenwald (2006)	**Common ground/Situation Awareness**	**1, 2, 3, 5, 6**
	Recognizing differences in the underlying meanings of shared symbols. (p. 6-9)	**1, 2, 5, 6**
	Sharing implications of information (p.11)	**1, 2, 3, 5, 6**
	Understanding the role of emotions in sharing information.	

Reference	Factor/Feature of information sharing	Activity theory component
	Re-establishing trust.	
Talja and Hansen (2006)	**Asynchronous activities**	**1, 2**
	Synchronous activities	**2, 5, 6**
	Co-located collaborations	**2, 4, 6**
	Remote collaborations	**1, 2, 3**
	Loosely coupled activities	**Knotworking**
	Loosely coupled activities	**Knotworking**
	Tightly coupled activities	**2, 6**
	Planned collaboration	**2, 3, 4**
	Unplanned collaboration	**Knotworking**
	Intragroup collaboration	**2, 5, 6**
	Intergroup collaboration	**2, 6**
	Direct collaboration	**2, 5, 6, Knotworking**
	Indirect collaboration	**1,2,5,6, Knotworking**
	Coordinated activities	**2, 4, 5, 6**
	Differentiated activities (p. 124)	**6**
Rafaeli and Raban (2005)	Behavioral factors	
	Social factors	
	Economic factors	
	Legal factors	
	Technological influences	
Talja (2002)	Person sources	
	Documentary sources	
	Formal channel	
	Informal channel	
	Technical information	
	Paradigmatic sharing	
	Directive sharing	
	Strategic sharing	**1, 2, 3, 4, 6**
	Paradigmatic sharing	**1, 2, 3, 4, 6**
	Directive sharing	**1, 2, 3, 4, 6**
	Social sharing	**1, 2, 3, 4, 6**
Cramton (2001)	Mutual knowledge problem:	
	failure to communicate and retain contextual information	
	difficulty in communicating and understanding the salience of information	
	differences in speed of access to information	
	difficulty with interpreting the meaning of silence (p.346)	**2, 5, 6**

Table 3-2: Mapping of information sharing factors into activity system dimensions

3.4.4 Other theories considered but not selected for this study

Evaluation of a set of alternative theoretical frameworks (Table 3-3) led to the conclusion that activity theory was the best fit for this study. Of these alternative theories considered, the most general is Maslow's Theory of Human Motivation (Maslow, 1943).

Maslow's Theory of Human Motivation (Maslow's Hierarchy)	Maslow (1943)
Grounded Theory	Glaser and Strauss (1967)
Sense-making	Dervin (1983, 1992, 1999)
Information Sharing Theory and Hidden Profile	Stasser and Titus (1985)
Theory of Information Sharing	Constant et al. (1994)
Information Acquiring and Sharing (IA&S)	Rioux (2004)
Grounded Theory in Study of Information Sharing	Razavi and Iverson (2006)

Table 3-3: Other theories considered for this study

The vantage point is a psychological perspective, and is the source of Maslow's Hierarchy of Needs, shown below in Figure 3-2. The relevant idea under consideration was that information sharing (or knowledge sharing) could be considered to be intrinsic and/or external incentives for knowledge workers (Hendriks, 1999) for the higher order levels of the needs hierarchy (Szirtes, 2011). However, this theory was almost too general (and at the same time too narrow) for this study. The over-generality is that it starts with the individual motivation and the internal psychological processes involved in the human condition It does not provide any framework to consider organizational or collaboration/process related factors, and would require augmentation. For that reason, activity theory is a more complete theoretical framework for investigating human information behaviors.

Many information behavior studies utilized grounded theory (Glaser & Strauss, 1967; Razavi & Iverson, 2006), a broad methodological theory particularly suited for exploration of new research areas. For instance, Razavi and Iverson's (2006) study provides a good example of a grounded theory study of information sharing in a knowledge-based, personal learning space, looking at privacy and trust issues. Grounded theory did not fit as well as activity theory because there was a rich literature foundation in many dimensions of information. In addition, activity theory was judged to be a stronger theoretical framework because it provided a structure for so many elements of importance in this study: the objective of the work, the

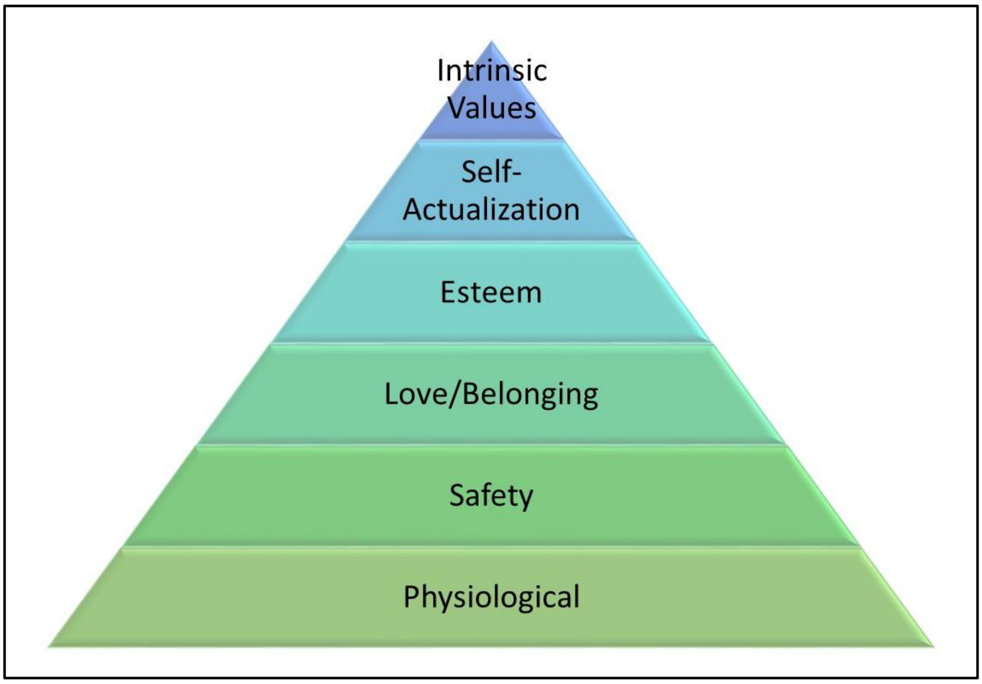


Figure 3-2: Maslow's hierarchy of needs (Skelsey, 2014, p.983)

community and rules, the contradictions which can lead to innovation or learning, and the dynamic “coming together” of people in knotworking.

Dervin’s (1983, 1992, 1999) sense-making theory was another one considered for this research, and was appealing for several reasons: the user-centered focus, the ability to extend the theoretical structure to multiple people in collective and organizational settings, the combined theoretical and methodological components, and the widespread adoption of sense making across many disciplines and studies. After examination, however, it was judged to be better suited to the activity of going out to obtain information, and not such a good fit for the activity of information sharing, a proactive activity to “push” information to another. The sense-making theory provides a constructivist, user centric approach to understanding how “people construct sense of their worlds and, in particular, how they construct information needs and uses for information in the process of sense-making” as they move through time and space (Dervin, 1983, p. 3). The idea of focusing on situational (work) contexts to understand the collaborative impact of information sharing is important for addressing the barriers to success of software projects discussed earlier.

There are three aspects of Dervin’s work relevant to the research questions, which activity theory shares. First, the contextual focus on addressing the changing environment around us to close a gap -- either as an individual in everyday life, or individually or collectively in a specialized work or social setting – is relevant. Dervin's conceptualization of "situations -- gaps -- uses" focuses attention on these facets. It was, however, awkward to frame information sharing in a subordinate role as a specialized sub-activity of information seeking.

Second, a fundamental aspect of the sense-making construct is situation awareness**.** A major missing affordance for virtual workgroups, who are collaborating but are not within physical proximity, is a mechanism of situational awareness which occurs naturally with co-located workgroups. Sense-making enables the exploration of this dimension of awareness of the overall context, from the viewpoint of the people themselves -- really the only perspective that matters. A user-centric focus is critical in order to adequately and accurately represent the user perspective. Activity theory provides situation awareness through the Community and Division of Labor dimensions.

Finally, the overlay of the sense-making theoretical constructs, with empirical data gained from micro-moment interviews directly from the participants, would have provided the opportunity to examine the use of information artifacts with the experience of the people in the sharing process. However, activity theory also provides this integrated capability with the Tools/Artifacts dimension, and some aspects of sense-making, particularly on a collective basis, can come into play in an activity-theory-based analysis.

The "Information Sharing Theory" of Stasser and Titus (1985), along with their further evolved theory called "Hidden Profile", is specific to information sharing activities and comes from the field of Psychology. They challenged the idea that group decision-making is more informed than individual decision-making, and examined the phenomenon of unevenly distributed information and the potential to "pool" it during discussion (Stasser & Titus, 2003). They also looked at the effect of new information disclosed during the group discussions and decision-making processes. The focus of Stasser and Titus' theories on decision-making was too narrow for this study, although they looked very promising initially. Activity theory can accommodate a much wider scope of project work elements, at a varying level of examination.

Constant et al. (1994) developed a theory of information sharing from their study of attitudes about information sharing:

> …attitudes about information sharing depend on the form of the information. Sharing tangible information work may depend on prosocial attitudes and norms of organizational ownership; sharing expertise may depend on people's own self-expressive needs. (p. 400)

As with Stasser and Titus, the focus on decision-making was too narrow in scope. In addition, it did not take into account any aspects of the overall "context" and constructs related to collaboration.

In Rioux's (2004, 2005) theoretical work of Information Acquiring and Sharing (IA&S), the framework defined is "sharing information found for others on the Web" (2004, p. 152). A conceptual model presents the cognitive state of the user, who encounters information on the web and remembers that someone they know would be interested in this information. This model was not selected due to different characteristics of this study, especially the evaluation of activities in a setting of variable size, and broader information channels (e.g., besides the web) in this study. For those reasons, activity theory provides a more comprehensive theoretical foundation for this study.

3.5 IMPLEMENTATION

This section describes the methodological implementation of data gathering and analysis in this study. Under the umbrella of activity theory, the researcher interviewed members of a workgroup (participants in this research) over an eight-month period using two repertory grid techniques: (a) elicitation, and (b) grid evaluation. These interviews were audio-recorded and transcripts produced. The relatively long period of phased interviewing provided time for the unfolding of contradictions, innovations, and other changes to occur. The researcher asked the participants to contribute example information artifacts and a few were collected. The data was analysed through activity system analysis and modeling (Engeström, 1999; Boer et al., 2002), guided by Mwanza's (2001) Eight-Step-Model and Kaptelinin, Nardi, and Macaulay's (1999) activity checklist. In addition, Leximancer 4 software analysis of the transcripts identified thematic trends and overall concepts across and within individual interviews.

3.5.1 The case as context: iProject

The researcher initially understood that the team (study participants) was a workgroup of approximately 20 members from one software project – including core workgroup members, people from affiliated projects, and stakeholders such as project managers. The people affiliated with this project, known as "iProject", are knowledge workers whose roles span research, technical design, architecture, implementation, and project management in the general space of data analytics environments, as well as services, software development, and products. Specific roles include:

- Researcher/Software Developer
- Researcher/Software Developer/Project Co-Lead
- Manager/Researcher
- Research Scientist
- Various Architect roles
- Various Chief Technical Officer roles
- Subject Matter Expert
- Technical Lead

Table 3-4 contains the full list of participant roles. This workgroup was comprised of people from multiple divisions and distributed geographic locations. It was also an important technology development project for both the Research and Product organizations, and had a high degree of personal commitment and investment by the workgroup members. The combination of these factors made it an excellent participant group.

3.5.2 Interviewing participants

One of the project co-leads suggested names of individuals to interview in November 2013, and all of these people granted interviews over an eight-month period between December 2013 and July 2014. New workgroup members and stakeholders who joined the project between December 2013 and July 2014 also participated in the interview process. In the end, there were 23 interviews. Organizationally and geographically, they represented multiple internal divisions within one company, and were physically located at multiple sites across the world (United States, India, China, and multiple locations in Europe). One or more stakeholders (project manager(s) and/or managers) associated with each direct participant participated in the interview process to bring in the stakeholder perspective from that organization. The full cohort participated in interviews, fulfilling the intention in the Research Design was to interview everyone in the iProject workgroup, both direct participants and management/project management stakeholders. This approach provided comprehensive reflections on information across the full project.

The interviews were conducted either face-to-face (if in the same location) or remotely (over the phone using a screen sharing tool). Each method worked well and had particular strengths. The interviews in the same location had the benefit of visual cues and body language, but the interviews conducted remotely captured more written content as the mechanism to create common ground in the interview. In face-to-face interviews, the interviewer was able to take private notes simultaneously; but in the remote interviews, the shared screen content constituted the notes.

The appropriate number of interviews to conduct in a qualitative study is a topic of much discussion and debate in the literature (Baker & Edwards, 2012). The total number of 23 in this study is a strong result because all of the core participants are included, as well as affiliated project stakeholders. In addition, extended workgroup members and stakeholders from affiliated projects are also included. This breadth, from individuals intensively engaged to those on the periphery, provides representation of varying engagement intensity across the workgroup effort. The study population also represents multidimensional distance. There is a mix of geographic locations, providing perspectives from multiple geographic vantage points, as well as across internal divisions of the company. In addition, there is a mix of roles performed for the project. Finally, there is diversity in the disciplinary backgrounds of the participants as reflected in their undergraduate, Masters, and PhD-level studies. Taken together, the multi-voiced interviews are of sufficient size to compare viewpoints on the same phenomena and events.

Table 3-4 shows the demographics of participants including (from left to right): (1) the job role they perform in the project affiliated with iProject, (2) their educational background, (3) categorization of their job role in the larger organization, (4) the company division they report to, (5) distance in miles from Location 1 (the location of the core team), (6) time difference from the core team in Location 1, (7) roughly for how long the participant has been working on the project affiliated with iProject, and (8) a unique participant code. It is important to note that job role in this case is their assigned job role in their employment, *not* an informal role adopted within the workgroup, such as a gatekeeper role, to help the workgroup function better.

(1) Role	**(2) Discipline**	**(3) Role category**	**(4)Unit**	**(5) Location offset**	**(6)Time Zone offset**	**(7) Project tenure**	**(8) Participant Code**
Researcher/ Software Developer	MS Computer Science; PhD in progress	Core iProject	R&D	Location 1	0	Beginning; ~2 years	RSD1
Researcher/ Software Developer	BS, PhD Computer Science	Core iProject	R&D	Location 1	0	Beginning; ~2 years	RSD2
Researcher/ Software Developer	BS, MS Computer Science/Mathematics	Core iProject	R&D	Location 1	0	Beginning; ~2 years	RSD3
Researcher/ Software Developer	BA Mathematics, PhD Computer Science	Core iProject	R&D	Location 1	0	early	RSD4
Researcher/ Software Developer	BS, MS Computer Science	Core iProject	R&D	Location 1	0	< 1 month	RSD5
Researcher/ Software Developer	BA Mathematics; PhD Computer Science	Core iProject	R&D	Location 1	0	middle	RSD6
Researcher/ Software Developer	Masters level, Computer Engineering; PhD in progress	Core iProject	R&D	Location 1	0	June-Aug, 2014	RSD7
Researcher, Chief Architect	Masters level, Mathematics	Affiliated Research	R&D	+ 2900 miles	+3 hours	over the past year	RCA
Researcher/ Software Developer	BA, MS, PhD Computer Science	Affiliated Research	R&D	+ 5900 miles	+15 hours	Summer, Fall 2013	RSD8
Researcher/ Manager	B.Tech, MS, PhD Computer Science	Affiliated Research	R&D	+8700 miles	+12.5 hours	Summer, Fall 2013	RM1
Research Scientist	M Tech, Communication Engineering, PhD Computer Science & Engineering	Affiliated Research	R&D	+8700 miles	+12.5 hours	Summer, Fall 2013	RS
Researcher/ Manager	BS, Computer Science; MS, Computer Science; MBA	Affiliated Research	R&D	+ 5900 miles	+15 hours	Summer, Fall 2013	RM2
Chief Architect	PhD, Computer Engineering	Product	Product	+ 2900 miles	+3 hours	2013	CA
Integration Architect	Masters, Computer Science	Product	Product	+5700 miles	+9 hours	2013	IA
Architect	BA, Computer Science	Product	Product	+7300 miles	+ 10 hours	Summer, Fall 2013	A2
CTO	PhD, Computer Science	Product	Product	+ 6 miles	0	ongoing - loosely coupled	CTO
Researcher/ Manager	AB Applied Mathematics, PhD Computer Science	Stakeholder iProject	R&D	Location 1	0	beginning; ~2 years	RM3
Researcher/ Manager	BA, Art and Anthropology	Stakeholder iProject	R&D	Location 1	0	1+ years	RM4
Technical Lead	BS, Mathematics, PhD Computer Science	Stakeholder iProject	R&D	Location 1	0	Beginning; ~2 years	TL

(1) Role	(2) Discipline	(3) Role category	(4)Unit	(5) Location offset	(6)Time Zone offset	(7) Project tenure	(8) Participant Code
Researcher/ Manager	BS, Computer Science & Honors Mathematics, MS Computer Science	Stakeholder iProject	R&D	Location 1	0	2013	RM4
Architect	AB, SM, PhD Mathematics	Stakeholder iProject	R&D	Location 1	0	2013	A2
Subject Matter Expert	BS, Clinical Laboratory science, MS Computer Science	Stakeholder iProject	R&D	Location 1	0	Summer-October, 2013	SME
Researcher/ Manager	MS, PhD, Cognitive Psychology	Affiliated Research	R&D	+ 2900 miles	+3 hours	none	RM5

Table 3-4: Demographics of the 23 study participants

3.5.3 Gathering data with the Repertory Grid Interview Technique

A two-phase repertory grid interview technique (RGT) was used to ensure that the voice and perspective of each participant be captured authentically, and to bring those unique viewpoints into the activity theory analysis, which is also strongly user-contextual. The two phases implemented were (1) elicitation of information sharing mechanisms, followed by (2) rating of a single repertory grid constructed by the researcher from phase one

Phase one consisted of 23 participants interviewed for at least one hour, sometimes in two sessions. Of these 23 participants, 22 were willing to contribute their examples of shared information mechanisms, followed by the qualitative characteristics of those mechanisms. (In one case, the interview reverted to a semi-structured discussion, in order to salvage some value from the interview. This was due to the participant's unresponsiveness to the repertory grid elicitation prompt.) Phase two – the evaluation of a single repertory grid by the participants -- started after all 23 interviews in phase one was completed. A single consolidated grid was constructed by the researcher by selecting common components – elements (what was shared), and constructs (descriptive adjectives about what was shared) from the 23 interviews completed in phase one. 22 participants received this grid electronically. (One participant left the company after the phase one interview and thus was not available to be included in phase two.) At the end, due to workload, time constraint, job change, and other factors, six participants did not complete or return their grids. Consequently, there were only 16 completed grids from phase two. The interviews of phase one, particularly the repertory grid elicitation segment of each interview, did produce a rich data corpus.

The researcher did not mention activity theory framework to the participants in the main study, a protocol change inspired by the first pilot. In the main study, the components of the activity theory framework emerged naturally and implicitly in the discussions of how information sharing occurred.

3.5.4 Implementation of the data gathering method

The prompt at the beginning of the interview was minimal. After explaining the purpose of the study and asking the participant for a brief statement on their educational background and their role on the iProject, the interviewer moved to the substance of the interview with the prompt “what is shared?” Each “thing” named was written down on an index card, or on a virtual index card in a shared virtual computer screen.

It was difficult to get some people to follow the RGT format; one person did not. If they did not, the interview proceeded as a semi-structured interview, with some additional prompts, about the work and the information exchanges in the project. The participants were able then to name information sharing mechanisms as they went along. A few people were comfortable right away and readily provided rich commentary on the sharing activity in their project.

In the next step, the participant discussed the associated characteristics of those mechanisms. This occurred in two ways. The first technique was through

Figure 3-3: Configuration of index cards after a face-to-face interview

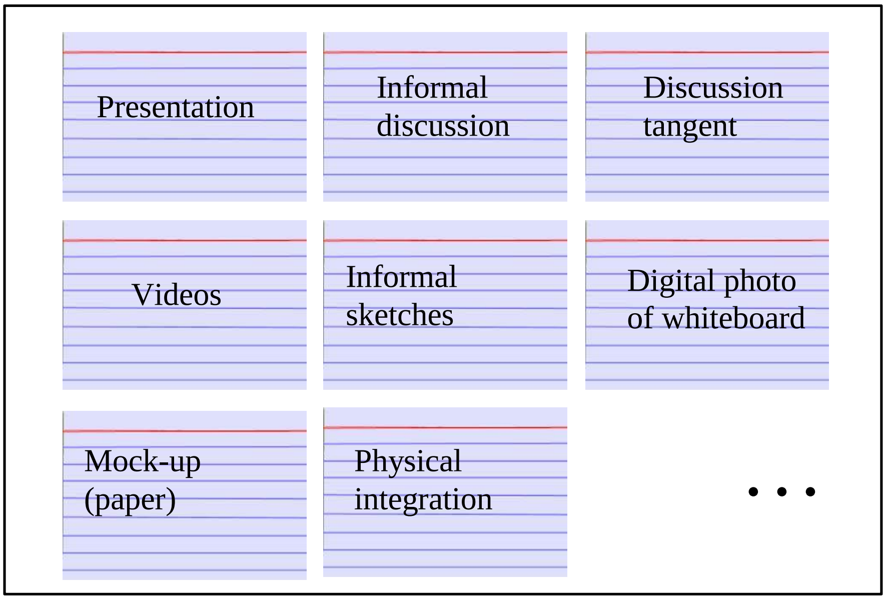


Figure 3-4: "Virtual note cards" created in interviews conducted remotely

participant narrative at the time they named the mechanisms, as it was natural for them to provide some description and explanation at that time. After the participant named all their information sharing mechanisms, the participant performed the repertory grid "construct" elicitation protocol successively for a varying number of times. This was a sequence of first selecting three mechanisms (written on index cards as shown in Figure 3-3, or on card images online as shown in Figure 3-4) and then grouping two of them together that are similar in some way, and contrasting those two with the third item. This technique enabled the subject to identify and describe sharing in the project completely in his or her own words and their own worldview.

It required self-control to minimize "helping" at the beginning of the interview as people got used to what they were to do. Many asked for answers, which would have defeated the intent of the protocol of participants answering in their own terminology and worldview. Similarly, with the successive grouping of three information sharing mechanisms at a time, and a comparison of similar and contrasting characteristics, there was a variety of comfort level with the process, and a varying number of constructs offered by the participants. Some participants were able to rapidly name many Elements and Constructs; however, some participants who named many Elements were able to name only a few associated Constructs. This repertory grid framework, unfamiliar to all participants, was a double-edged sword: difficult to get the elicitation started, but opening a fresh entry into the experiences of the participants because it was different. The grid rating activity had similar awkwardness, as many participants reached out for guidance after receiving the email with the grid and the associated request to complete it.

The researcher asked the participants to provide examples of shared information artifacts (e.g., documents, email, and project information stored in data repositories). These materials served as an additional source of information supplementary to the interview transcripts. These included several documents and wiki URLs.

3.5.5 Unit of analysis: activity

The unit of analysis in activity theory is the human work activity itself (Engeström et al., 1999). This is a great strength in analysis. The comprehensive analytical viewpoint brings multiple advantages in the study of a complex, multi-person, and technology-mediated phenomenon:

1. While simultaneously maintaining a perspective of the full “landscape”, analysis at multiple levels of granularity -- from the highest and most general level of the system with aggregated participants, all the way down to the perspective of a single participant, and mid-levels in between – provides a rich analysis.
2. The dimension of “subject” provides the capability for multi-voice analysis, and comparison across multiple dimensions.
3. The integrity of the human activity is preserved, and not overly simplified.
4. Analyses of changes over time (longitudinal analysis) provide a viewpoint of the evolution of the activity system.

The research design of this study takes advantage of all of these strengths, and benefits from the analytical insights of this approach.

3.5.6 Data Analysis: Overview

The sequence of analysis preparation of the interviews included recording, transcription, and correction. The sequence of analysis of the interviews included decomposition, organization, and loading into content analysis tools. Extensive data analysis was done of the recordings, transcripts, lists of information sharing mechanisms, and the described characteristics of those mechanisms. A complete mapping of the information sharing mechanisms using a spreadsheet tool provided a taxonomic visualization; a second spreadsheet constructed by mining the corrected transcripts and mapping the activity system components provided an intra- and inter-interview visualization, and the Leximancer 4 tool was utilized with both corrected transcripts and spreadsheet output to identify important concepts and their relationships. The visualization and machine learning capability of this tool provided another analytical view on the interview narratives and derivative data. The purpose of the analysis was to gain a thematic understanding of the full narratives, to cluster and contrast the narrative text by demographics and job role categorizations, and to explore the results from a different angle. Also important for the researcher was an immersion in the narratives and worldviews of each participant, and a detailed familiarity with the content.

The combination of these three analytical activities enabled method triangulation between analysis of the elicited repertory grid elements and constructs, the activity systems, and the automated conceptual and thematic analysis utilizing a machine-learning tool. As to the data collected from phase two, since the number of completed grids was incomplete, there was only a limited manual review but not a full analysis. Activity system representations at key points of the project and at varying levels of granularity provided a nuanced view of the information sharing and collaborative dynamics. Table 3-5 provides a summary of the data analysis implementation, detailing the steps, the input data for each step, derivative data produced, the relevant literature related to this method. The following sections provide additional descriptions about this analysis.

3.5.7 Data analysis: data cleansing and preparation

Review and analysis of the 23 interview transcripts first occurred prior to any corrections. However, it became evident while working with the uncorrected transcripts that clean-up and correction was required in order to maximize the fidelity and the usefulness of the information provided by the participants. This painstaking interactive process occurred by playing each recording while stepping through the corresponding transcript. This provided a great and unexpected benefit: the process of this correction provided the opportunity for deep immersion into the content of the interviews over an intense, focused period, jump-starting a deeper thought process about the information sharing experiences of the participants.

Next, a list of information sharing mechanisms by interview was constructed and categorized in a spreadsheet, providing a taxonomy of information sharing mechanisms within each interview, and across all interviews. This provided an organized and consolidated view of the data. For example, there were many variations of "meeting", some identical, some similar, and some very different across the interviews. A spreadsheet enabled a consolidated view, along with descriptions of those elements.

3.5.8 Data analysis: activity system analysis and modeling

In a way, Table 3-5 is also a summary of the structured and unstructured interview data that served as the basis for activity system analysis and analysis/modeling of activity system components. A second spreadsheet was created with interview information and quotes for each participant for each of the nine vertices in the activity system model diagram. This "deconstructed activity system" data enabled a conceptualized view across each interview on all the activity system dimensions, across interviews, and across participants. This technique enabled visualization and comparison of differences across individual activity systems, as well as visualization of higher-level activity systems.

Activity system models show a situation at varying levels of detail, at varying points in time, for an individual activity system, networked activity systems, and the overall activity system. An example of such model diagram is included here (Figure 3-5) for illustrative purposes. Review and analysis of the resulting activity system information provided the capability to see variations and contradictions. Identification of areas of activity system development ("breaking away") and loosely coupled initiatives ("knotworking"), and the relationship of these areas to the process of information sharing that occurred are highlighted in the Findings Chapter (Chapter 4).

Next, the transcripts, the identified information sharing mechanisms, and their associated characteristics were utilized as source data for more analysis, following the analytical process defined by Boer et al. (2002, p. 1491), according to the activity theory system construct. This analytical process – originally designed for

Input data for analysis	Derivative data produced	Analysis method
• Interview recordings • Interview transcripts • Interview notes	• Corrected transcripts • List of information sharing mechanisms • Description of each information sharing mechanism characteristics	• Data preparation and correction
• Corrected transcripts • List of information sharing mechanisms (elicited repertory grid elements) • Description of each information sharing mechanism characteristics (elicited repertory grid constructs)	• Taxonomy spreadsheet of categorized information sharing mechanisms • Single repertory grid	• Manual data categorization and analysis
• Corrected transcripts • List of information sharing mechanisms (elicited repertory grid elements) • Description of each information sharing mechanisms' characteristics (elicited repertory grid constructs) • Taxonomy spreadsheet of categorized information sharing mechanisms	• Activity system component spreadsheet -- identification of each of the 6 activity system components per interview, in narrative text block.	• Activity system analysis and modeling (Engeström et al., 1999); and guided by Boer et al.'s (2002, p. 1491) activity theory approach, Mwanza's Eight-step-Model (2001), and Kaptelinin et al.'s (1999) activity checklist.)
• Corrected transcripts • Initial activity system model diagrams	• Decomposition of activity systems -- network of activity system model diagrams	• Activity system analysis and modeling (Engeström et al., 1999); and guided by Boer et al. (2002, p. 1491) – step 2, Mwanza's Eight-Step-Model (2001), and Kaptelinin et al.'s (1999) activity checklist.)
• Decomposition of activity systems: a network of activity system model diagrams	• Selection of activity systems for detailed analysis. Detailed evaluation of mediating processes in/between the 6 components of each activity system (Boer et al., 2002, p.1491)	• Activity system analysis and modeling (Engeström et al., 1999); guided by Boer et al. (2002, p.1491) – step 3, Mwanza's Eight-Step-Model (2001); and Kaptelinin et al.'s (1999) activity checklist.)
• Detailed evaluation of mediating processes between the 6 components of each selected activity system Boer et al., 2002, p.1491)	• Analysis of how information sharing is manifested within and between activity systems	• Activity system analysis and modeling (Engeström et al., 1999); guided by Boer et al. (2002, p.1491) – step 4, Mwanza's Eight-Step-Model (2001), and Kaptelinin et al.'s (1999) activity checklist.)

Input data for analysis	Derivative data produced	Analysis method
• Shared documents provided by participants	• Document analysis	• Activity system analysis and modeling, guided by mediating artifacts hierarchy (Collins et al., 2002)
• Completed repertory grids with evaluations	• Observations	• Manual review using knowledge of the overall data collected.
• Corrected transcripts • Spreadsheet of categorized activity system components (interview quotes) • Spreadsheet of categorized information sharing mechanisms	• Leximancer 4 visualizations and text analytics output	• Conceptual and thematic analysis using Leximancer 4.

Table 3-5: Overview of data analysis

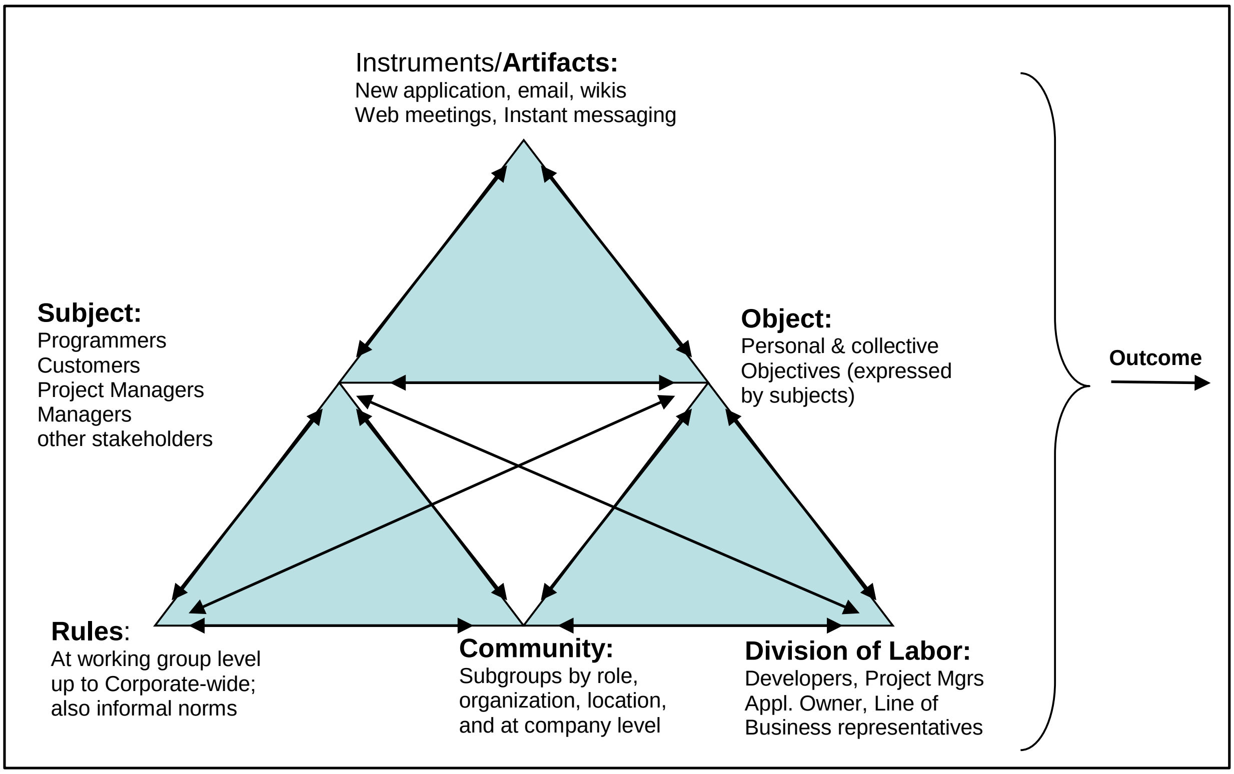


Figure 3-5: Activity system model example

studying knowledge sharing and adapted in this study for analysing information sharing -- focused at varying levels of granularity, going from individual elements up to the broadest activity system that can be drawn. This method consists of five steps, and Table 3-6 summarizes the adapted version.

1. Choose the organizational setting within information sharing to study and translate this organizational setting into an activity system (see paragraph 3).
2. Define activity systems at other contextual levels of analysis in order to zoom in or zoom out, till a level of abstraction is found which addresses the intended issues for understanding information sharing (see paragraph 4).
3. Describe the mediating processes between the components of each activity system by indicating the development of each component and the (potential) tensions within and between these components. Specify how the different activity systems are interrelating (see paragraphs 3 and 5).
4. Explore how information sharing reveals itself within and between the activity systems by relating it to the transformations of their objects and to existing or potential tensions (see paragraph 2 and 5).
5. Repeat the previous steps by taking the perspective of different subjects and reconcile the different perspectives. Relate the findings to the original activity system.

Table 3-6: Activity theory data analysis method (Boer et al., 2002, p.1491)

Activity System analysis – iProject level

Component	Jan-Feb, 2013	May-Oct, 2013	Feb, 2014	May-July, 2014
Tools/ Artifacts	Meetings, email, mockup, slides, presentation.	Meetings, email, mockups, slides, presentation, prototype code (real running system), milestone, poster	Meetings, email, prototype code, shared code repository	Meetings, email, prototype code, shared code repository
Subjects	iProject core developers (matrixed)	iProject core developers (matrixed), Subject Matter Expert, Product Architects, Affiliates	iProject core developers (consolidated), Product Architects, Affiliate	iProject core developers (consolidated), Product Architects, Affiliate
Object	Demonstrate project concepts and build support for project, future funding	Demonstrate project concepts at a conference about data analytics tooling, and obtain customer feedback. Continue to build support for project, future funding.	Shared: To deliver new data analytics capabilities into our product set, that can then be used to support some of the higher level goals for our department, and the new partnership mission, as well as our external customers.	Shared: To deliver new data analytics capabilities into our product set, that can then be used to support some of the higher level goals for our department, and the new partnership mission, as well as our external customers.
Rules	Conform to time slot, give positive impression.	Meet deadline, high quality customer interactions at event.	Balance commitments across missions	Secure internal income stream; balance commitments
Community	Analytics community, Research, Researchers/SW Developers, Stakeholders	Analytics community, Research, Product Division, external customers	Analytics community, Research, Product Division	Analytics community, Research, Product Division
Division of Labor	Specific team member created mockup, others participated in design discussions, volunteers	Core team implemented different components; Global Labs ran their own code on common dataset and contributed intermediate files; Product group created extension paradigm.	More core workgroup members added in reorg; ongoing collaboration with product group and affiliates	Continuation of core workgroup efforts and ongoing collaboration with product group and affiliates

Figure 3-6: Example of activity system analysis and components – project level

Activity System analysis – composite individual person

Component	Researcher/Software Developer
Tools/ Artifacts	**Telephone meeting, chat, screen-sharing, presentation slides, questions and answers via email.** On what to do for the ABC demo work: "I obtained this information from a con call, a telephone meeting, and by email." p3 "[We] used some desktop sharing systems such as – let me check - such as an instant meeting room in Lotus Notes and we have also another system that is called Cloud meeting. Yes. We use these two systems to share our desktop and to show, for example, slides of some results from the software output. At the same time, we will introduce a -- how do you say -- by telephone." p4 "For example, we usually ask some questions and provide some answers by notes, by email. For example, last year, there is an email that is send by [Person X] -- another software engineer -- and he asked what is the result of one software module and why there are some errors in putting the result into the system - into the Information Server system and I reply this email by introduce or by provide correct result from our software because the result provided by him is not the up to date one so I provide new generated result to him and finally, he imported the new generated result into the system correctly." p5 ". I didn't use a source control system because physically, we are loosely coupled with each other." p5 "He used one file to import into the system but that file is not the one that should be imported into the system. That - the file he used is intermediate result and it should not - yes. It should not be and they're not able to be imported the system so I provide him one that could be imported in the system and he finally, successfully do it." p6 Architecture not necessary: "Maybe - Architecture is not very necessary maybe but the input and output will be should be presented I think. And if the function - if the introduction is given, people will have a good understanding." p10 Instant Messaging: "Instant messaging is more efficient but usually we are located in different time zones maybe the other team member will not be online at same time. Yes and so I only could be - I only able to send email. Another thing is - another thing is if the question - it's not a quick one. I mean, there are some context I need to explain, I think I will preferably choose the email because I could write down many of the context of the questions and yes and the email would be better." p12
Subject	**Researcher/Software Developer, Remote Lab. "**Basically, I joined the project as a software engineer." p1
Object	"I think maybe to provide the functionality, the feature into the iProject and which will be demonstrated at the ABC Conference." p1
Rules	"What, if you do this, politically, how is it, you know, inside our lab, who do you take instruction from, how do you spend your time, what's going to work, what isn't going to work." p22
Community	On closer relationship with Lab#1 and Lab#2: "We're working on the same sorts of problems so obviously we will be more inclined to work together." p9
Division of Labor	It's a relatively small number of people that really drive it. It's probably a larger kind of ecosystem around it, the people that take an interest, people that occasionally step in and out, but if you look at the core drivers - there are probably a handful of people and then in that sense, I think amongst them there's probably clarity about what happens. You know, you don't need this kind of formalization that you would need if you had more geographically distributed people, if you had a larger team, if you had a situation where people just drop out for a few weeks and then come back again. And then you need to be much more formal and document everything." p11
Outcome	"To provide some new functionality for the product and so to provide some new features into the other product. Then there are three views and the first is curation and the second is enrichment and third is agile exploration. Basically the aim is to provide some functionality to explore or curate very large numbers of data. For example, for CSV file, for JSON files, and find some relationships between these various kinds and forms of more data." pp2-3 "Sorry. Because - actually, I didn't come to the ABC meeting. ABC

Figure 3-7 Example of activity system analysis and components – individual level

Figure 3-6 and 3-7 show examples of activity system analysis using data from the interview transcripts and the elicited information sharing tools/artifacts. Figure 3-7 is a composite constructed from multiple participants in order to protect participant anonymity. Snippets of the transcripts were included in order to provide richer detail and deeper context. Chapter 4 (Findings) contains the actual findings across all interviews, but these figures are included to provide a view of what was created using the interview data.

Kaptelinin et al.'s (1999) activity system checklist provided guidance to examine the activity theory components. The purpose of this checklist was to design or evaluate a user interface, and was adapted for this study of information sharing. The Activity Checklist provides a framework for exploring the five basic principles of activity theory, to consider as an integrated system:

1. Object-orientedness: "Every activity is directed toward something that objectively exists in the world, that is, an object" (p.28).
2. Hierarchical Structure of activity: Activities, Actions, Operations (p.29).
3. Internalization and Externalization: "activity theory emphasizes that internal activities cannot be understood if they are analyzed

separately, in isolation from external activities, because it is the constant transformation between external and internal that is the very basis of human cognition and activity" (p.29)

4. Mediation: With "emphasis on social factors and on the interaction between and their environments" (p.31).
5. Development: "Activity theory requires that human interaction with reality be analyzed in the context of development" (p.32).

The Activity checklist (Kaptelinin et al., 1999) consists of general questions in the following four areas:

- Means and ends -- the extent to which the technology facilitates and constrains the attainment of users' goals and the impact of the technology on provoking or resolving conflicts between different goals.
- Social and physical aspects of the environment – integration of target technology with requirements, tools, resources, and social rules of the environment.
- Learning, cognition, and articulation – internal versus external components of activity and support of their mutual transformations with target technology.
- Development – developmental transformation of the foregoing components as a whole. (p. 33)

To make the checklist suitable for this research, the term "information sharing" replaced "technology" in the Activity checklist.

3.5.9 Data analysis: Leximancer 4 concept and thematic analysis

Phase one of the repertory grid interview was intensive in nature and yielded rich data on information sharing mechanisms and descriptive narratives about those mechanisms. Unfortunately, phase two – evaluation of a single repertory grid based on consolidated input across the phase one interviews – had incomplete results. Due to participant time constraints and business pressures, six of the 22 phase two participants did not complete the grid evaluation, leading to an incomplete data set. Brief manual analysis of the 16 completed grids enhanced understanding of the interview data. Use of Leximancer with the full transcripts provided additional and complementary analysis, since the originally planned analysis of the grids was not completed. This analysis enabled enrichment of the findings from the activity systems analysis of phase one data. Thus, the final analysis represents multiple analysis methods on the elicited constructs and elements in the phase one repertory grid interview structure, and the interview narratives. These tend to converge to provide the necessary breadth and depth of participant perspectives on information sharing in distance collaboration from both the individual's perspective and its place and importance within the group.

Leximancer 4 software enabled analysis of the interview transcripts for themes, concepts, and connections with natural language and machine learning algorithms. Thematic clustering across the interviews provided a perspective on both unique and common elements across the interviews. Relevant demographic characteristics for each participant, including disciplinary and educational backgrounds, work experience; role on the project, organizational position, and geographic work location provided a multi-dimensional perspective. These categorical data (and other characteristics in the study population as well) were useful for identifying differences within and across thematic dimensions. No quantitative analysis was performed on the repertory grids since an insufficient number were completed.

3.5.10 Data analysis: shared information examples

The analysis of shared information examples provided by the participants in multiple ways was not completed. The number of such artifacts obtained from the participants was so small that it would not be meaningful to proceed as planned. Instead, a manual analysis of the shared information artifacts provided by any given participant and comparison to the participant's interview narrative was done.

3.5.11 Pilot studies

Prior to the main study, two pilot studies to test and refine the interview protocol – both the data gathering and data analysis processes – were completed. The first pilot study, conducted in April 2011, consisted of interviewing two members of a software engineering workgroup utilizing the semi-structured activity system interview, and reviewing documents provided by the participants. The second pilot study, conducted in April 2013, consisted of interviewing two different individuals utilizing the repertory grid interview technique.

Pilot Study One

Problems/objectives of pilot study one

The purpose of the pilot study was to ensure that the semi-structured interview protocol worked well with the participants, to identify any needed modifications, and to check the fit of the data to activity theory.

Process of implementation

The researcher invited two members, randomly selected from a company workgroup, to participate in the pilot study via an interview estimated to last for one hour in duration. The researcher provided the Recruitment flyer and the Informed Consent form from the QUT Ethics submission. Both agreed to participate, and a specific date and time were established. The AT&T telephone conference number enabled recording, as well as a web conference system. The two recorders enabled redundancy in case one of the recordings failed.

Subjects/Participants

The first subject was a technical lead from the Research workgroup who had worked for the company for 13 years and had been on the project for 4.5 years. His/her educational background included BS in Computer Science and MS in Software Engineering Management. The other subject was a Quality Assurance (Test) Management manager with 15 years tenure at the company, and a BS degree in Computer Science and Mathematics. Both participants were located in the United States, but in different cities, and were separate from other members of the project workgroup.

Data collection and content analysis

The interviews were recorded using two recording methods (for most interviews) in case of failure. AT&T Conferencing produced a transcript, subsequently anonymized for purposes of confidentiality, with names and other sensitive specifics removed. Comparison of the transcript with a recording enabled corrections, followed by anonymization. Leximancer 4 provided content analysis on the collected information artifacts and comparison with the interview narratives.

Preliminary analysis

The deeply situated contextual perspective of each person came through very strongly in all aspects of the interview. It was clear that they were talking about the same project, but there were significant differences and emphases reflected in the narratives given about the project characteristics, roles, objectives, and other characteristics. As expected, this resulted in a unique activity system for each person that has some similarities, but also some significant variations. Following are the highlights, augmented by the two activity system model diagrams.

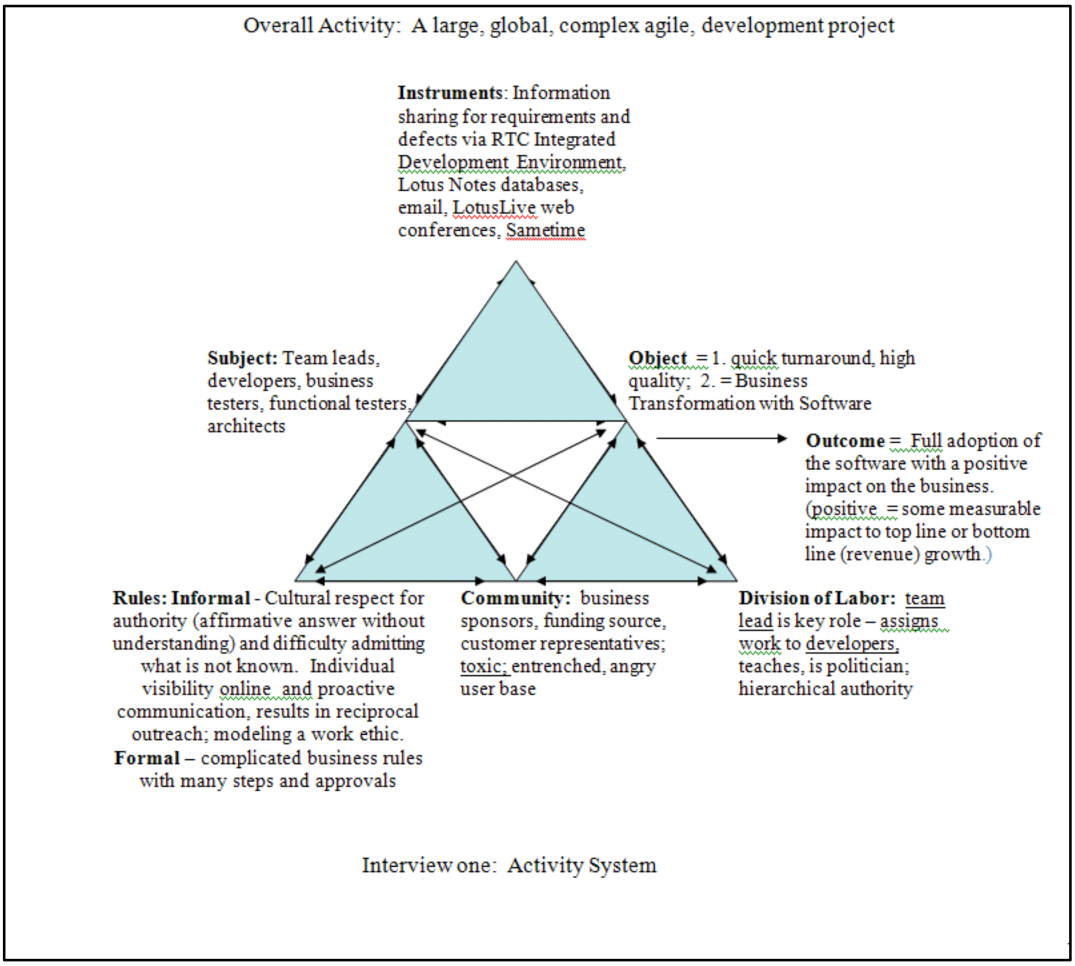


Figure 3-8: Activity system for pilot one, interview #1

Figure 3-8 shows the activity system as perceived and described by participant #1, constructed from the semi-structured interview narratives. Of particular note is the central role of the workgroup lead, the position held by the participant. This participant stressed the critical role of the (Development) workgroup leads, and the development workgroup, throughout the interview. Also important to note is the characterization of the community: "the business sponsors … our funding source, and then, you know, customer representatives. And I would classify our work environment as toxic".

Figure 3-9 details the activity system described by participant #2. Again, demonstration of the contextual perspective is strong, this time emphasizing the Quality Assurance (QA) activity.

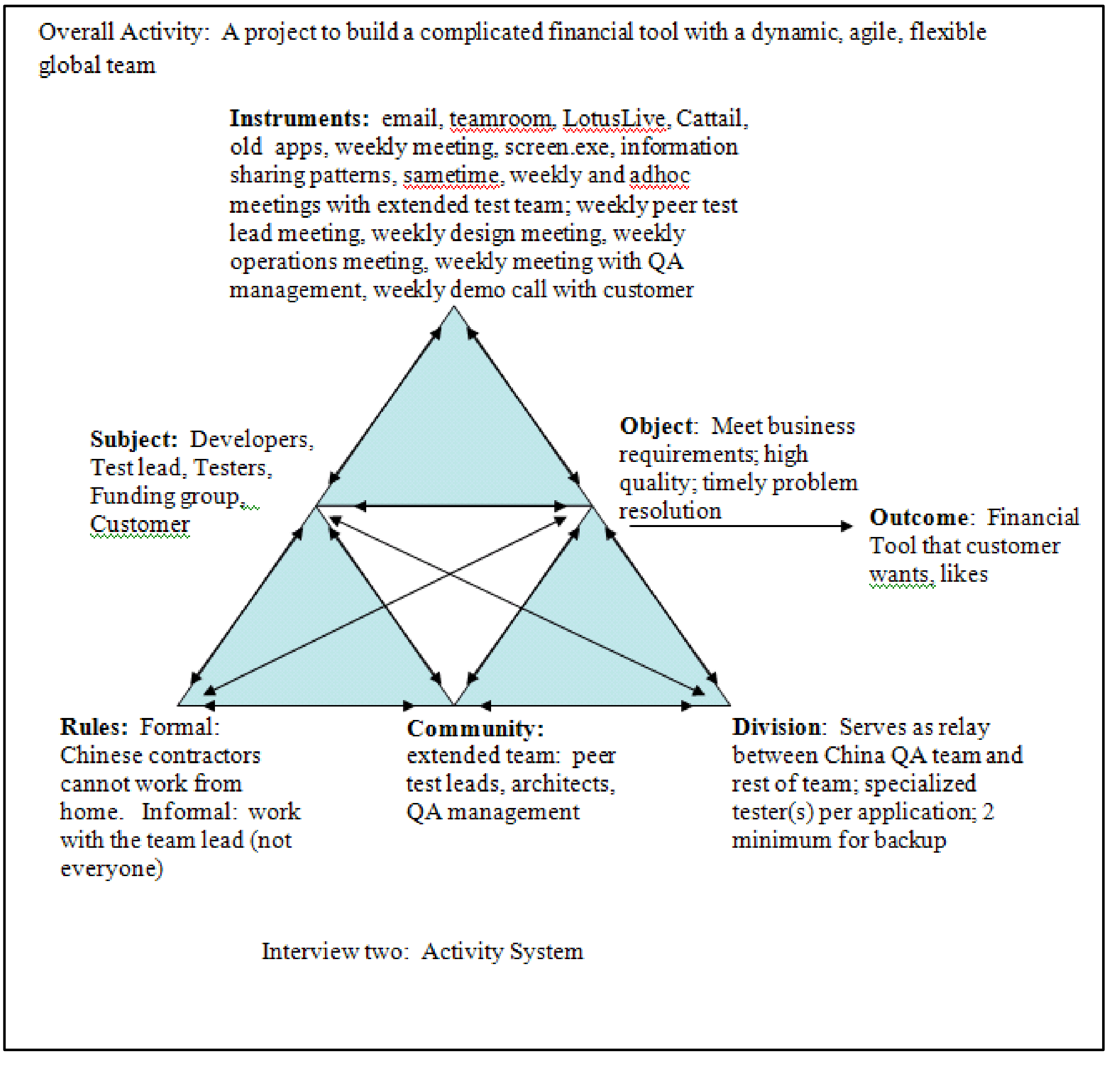


Figure 3-9: Activity system for pilot one, interview #2

Many individual adjustments to address collaboration issues, some related to information sharing, were noted. These included tactics such as adding an explicit confirmation of understanding in discussions in order to give permission or to admit lack of knowledge and/or understanding, extending work hours to overlap with the normal work hours of the extended workgroup, making proactive efforts to be available via instant messaging for dynamic contact, and actively reaching out to workgroup members in other time zones, and trying to achieve "constant communication".

The very open environment of sharing information seemed to result in some tension within the workgroup, leading one participant to suggest filtering of communication and limiting the ability to share information so openly. A notable observation is that the transparent, flat information environment contrasts with the hierarchical, sometimes politically charged, organizational environment. In addition, this very open environment of sharing information can result in confusion due to contradictions that emerge over time, as well as changes in requirements that occur. Each participant commented on the fast work pace. Finally, there are practices of conventions and patterns for information sharing by project participants, but not necessarily across the full workgroup, leading to potential misunderstandings within the workgroup.

Pilot Study Two

Although the interview was productive and relevant insights emerged from the interview protocol of the first pilot study, the interaction itself was awkward. The researcher added a second pilot study, with the purpose of improving the interview protocol and testing a different technique with less interviewer influence. The second pilot study, conducted in April 2013, consisted of interviewing two new individuals utilizing the repertory grid interview technique.

Problems/objectives of the pilot study

In the first pilot study, the use of the semi-structured activity system interview was awkward due to the direct focus on the constructs of activity theory (e.g., tools, object, rules), and the requirement for the interviewer to drive the interview on these topics. Between the first and second pilot studies, the researcher identified the repertory grid interview technique as a possible interview approach that would allow the participant's viewpoint to come through more naturally. The second pilot study was to ensure that the new protocol with repertory grid would work better, to identify any additional modifications needed, and to check the fit of the data to activity theory for analysis.

Process of implementation

The researcher randomly selected two new workgroup members from an existing software engineering project, and invited them to participate in an interview for approximately 1 hour in duration, with a follow-up conversation (estimated to run for about 30 minutes) to complete the ratings on the repertory grid. The researcher provided potential participants the Recruitment flyer and the Informed Consent form from the QUT Ethics submission. Both accepted the invitation, and a specific date and time were established.

Subjects/Participants

The first subject was a technical lead from the Research workgroup who had worked in the company for 5-10 years and had been on the project for several years. His/her educational background included an undergraduate degree in Computer Science and an MBA. The other subject was a Software Engineer who had also worked in the company for 5-10 years, held a BS degree in Computer Science and an MBA. Both participants were physically located in South America.

Data collection and content analysis

Two modes of interview recording via a telephone conference number as well as iPhone voice recording provided redundancy in case one of the recordings failed. AT&T Conference Service produced a transcript, which the researcher corrected by comparing the transcript against a recording. The researcher subsequently anonymized the transcript to ensure anonymity and confidentiality, with any identifying names and other sensitive specifics removed. Leximancer 4 provided content analysis on the collected information artifacts, and a comparison with interview narratives.

Preliminary analysis

Elicitation of the Elements (information sharing mechanisms) and Constructs (characteristics) via the repertory grid interview technique worked very well, once the participants became comfortable (in the first interview) that they would need to name the mechanisms of sharing, and associated characteristics, and that the context would not be provided by the interviewer. This technique was unfamiliar to both individuals, and required finesse on the part of the interviewer not to provide too much information to get the conversation started. Once it started, it went very smoothly. This also happened with rating the resulting grid, as they did not know how to complete the grid. The interviewer (researcher) provided minimal prompts/answers, such as "the scale of 1 to 5 is like a Likert scale, just with customized meanings on the two ends instead of something like "very dissatisfied" to "very satisfied".

The elicitation process for naming of information sharing mechanisms (repertory grid elements) provided, first, the mediation tool itself, and second, an opening for a description of context, which provided the seeds of Object, Community, and Division of Labor components. The selection of three elements (two grouped together for similarity and one contrasting) also provided more context and a view into how the participant thought about those components. Moreover, the subsequent rating in the grid provided a deeper view about how those elements compared and contrasted with each other, over all the defined constructs.

Table 3-7 shows the resulting grid for one of the interviews, with the mediating tools in bold, flanked by the contrasting poles with descriptions of the characteristics elicited through the triad groupings. The columns 1 and 5 represent contrasting poles of characteristics, similar to a Likert scale, but with contrasting pairs of adjective descriptors at each pole for each row. The participant rated each column noun in the context of each row of descriptors. As an example, there is a polarized view of the tools by the participant looking at the aspect of "real-time" vs. "asynchronous". Only "Cloud meetings to share screens" and "Phone call" have a real-time quality, as the others (rated at '5') are all highly asynchronous. It is interesting that there is nothing in-between.

Moving to the activity system, Figure 3-10 shows a high-level activity system for one of the interviews. With the narrative description provided in the interview (transcript), it is possible to zoom in to levels that are more granular, and create new activity system views for a particular described situation.

Two changes to the main study protocol came from learnings in the second pilot study. Specifically,

- the decision to create a single repertory grid toward the end of the interview process from information gathered from the interviews. One option was to have participants rate a grid created from their individual data, and/or one constructed from a consolidation of data across the participants. Due to participants' time constraints, the researcher asked them to complete a common, consolidated grid.

- for sensitivity to minimal prompts from the interviewer, the initial prompt of "what information do you share?" was changed to "what do you share?"

As in the first pilot study, the contextual view of the participants came through, but through a more natural interview exchange. It provided the raw elements needed for the activity system analysis and modeling, as well as deep information about how people characterize what is shared, which is of greatest interest in this research.

1	**Power-points**	**Tech arch**	**Mock-ups**	**Cloud meetings to share screens**	**Req'ts doc**	**Phone call**	**e-mails**	**Proof of concept**	**Object model**	**5**
Real-time	5	5	5	1	5	1	5	5	5	Asynch.
General purpose (comm.)	1	5	4	1	5	1	1	4	5	Specific purpose
Static	1	1	1	5	1	5	5	2	1	Flexible
Used to comm. visually	1	1	1	1	1	5	1	1	1	Voice only; constrained
Used at development time	3	2	4	3	4	3	3	4	2	Used at beginning of project
Used to disambiguate, validate understanding & capture ideas, make abstract or ambiguous ideas concrete	3	4	1	3	5	3	3	1	3	Convey decisions that are already taken
Target audience: customer	3	5	2	3	4	3	3	3	5	Target audience: technical team
Used with customer early in the project	3	5	3	1	3	1	1	3	5	Voice only; constrained

Table 3-7: Repertory grid from interview #1 of the second pilot study

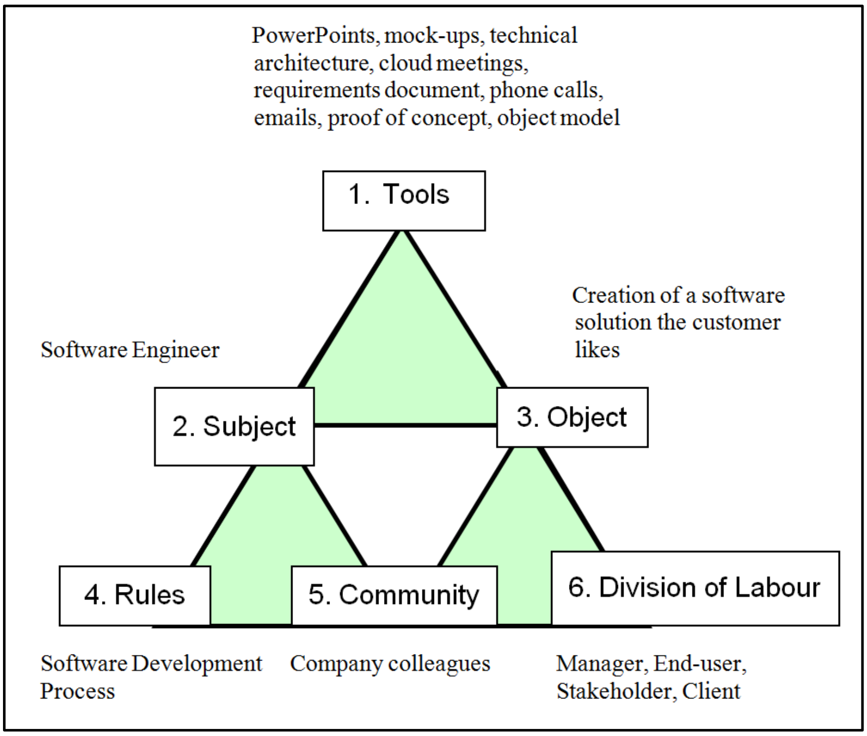


Figure 3-10: Example activity system from the second pilot study

3.6 ETHICS CLEARANCE INFORMATION AND RESARCH SUPPORT

The Chair of QUT University Human Research Ethics Committee approved application 1100000030 ("Collaborative information sharing in virtual teams") on March 3, 2011 in the category "Human Low Risk". The Ethics Committee subsequently approved a variation on February 13, 2013 for the addition of the semi-structured interview, which includes use of the repertory grid interview technique to compare different kinds of information that is shared, and the participant informed-consent form was amended. An extension by the Ethics Committee to March 3, 2015 was granted on March 5, 2014, and later extended to March 3, 2016. This research project was supported by the International Business Machines Corporation (IBM) through their support of the PhD program of the researcher.

3.7 SUMMARY

The opportunity to study a complex, technically intensive human activity in a real-life setting is a privilege, and activity theory in combination with a semi-structured repertory grid interview provided a powerful approach to build understanding in this area. This chapter describes the overall research design and implementation for this information sharing study in a context of distance collaboration and agile software engineering. Critical learning occurred in the pilot studies about how to establish an interview environment with the participant, resulting in a receptive and open atmosphere. The steps taken, from using a minimal prompt, to the use of the repertory grid elicitation technique, resulted in the capture of the user voice and perspective. Moreover, despite not being able to gather some planned data, the gathered data had richness, depth and substance. It was a lesson in sensitivity to the regular work commitments of the participants and respect for their time. The analysis which preserved that precious and contextual data was also effective in yielding multiple interesting insights.

Chapter 4: Findings

4.1 CHAPTER OVERVIEW

Now we look at what we learned from the data about people's information sharing and collaboration. We enter the individual and collective world of the participants and examine the information sharing mechanisms – both events and artifacts.

- First, a consolidated view of the information sharing mechanisms across the 23 unique contexts of the participants is provided to give a sense of the landscape, followed by additional detail about information sharing activities, events, and artifacts identified in the study.
- A discussion of the experience of information sharing follows, reviewing the relationship of information sharing to information seeking, collaboration activities with information sharing, and degrees of information sharing.
- Distance in information sharing is discussed next, with a focus on context, manifestation, emergence of new ideas, and the dynamic and flexible relationships of multiple activity systems (Knotworking) as demonstrated in the relationships between peer projects of iProject.
- The activity theory perspective comes next, as activity systems of varying granularities are next discussed, followed by a look at the longitudinal development of the highest level iProject activity system over 19 months. Identified contradictions present at varying times are reviewed, and innovations that could be attributed to activity system changes in the "Zone of Proximal Development" are presented.

To guide the discussion of findings as well as to provide a context, the research problem and questions are re-stated here, as a reminder. The core question of this research is: "How does information sharing occur in the distance collaboration of virtual teams?" The research problem centers on the phenomenon of information sharing in workgroups, and the associated challenges to collaboration and the accomplishment of work objectives by an extended project team. The subordinate research questions for the study are:

1. How do information sharing activities manifest themselves in distance collaboration?
2. When and in what kinds of circumstances does information sharing occur in distance collaboration?
3. What types of information sharing behaviors and forms of shared information can be identified?
4. What attributes are related to different types and forms of information sharing in distance collaboration?
5. What purposes does information sharing serve in distance collaboration?

4.2 INFORMATION SHARING: THE LANDSCAPE

4.2.1 Summary of the information sharing mechanisms

For this discussion, the term "information sharing mechanism" is an umbrella term for the answers by the participants to "what was shared". These are nouns – entities that identify and describe something shared. There are two sub-categories of information sharing mechanisms. The first is event, an interactive session of varying composition in which the sharing of information occurs, attended either in person or remotely. Examples of information sharing events are meetings and demos of a software system or application. The second is artifact, a human-constructed thing exchanged between people and utilized independently and asynchronously from the sender. Information sharing mechanisms, events, and artifacts all belong to the Tools/Artifacts category of an activity system.

As explained in the previous chapter, 22 of the 23 subjects identified mechanisms of information sharing (named "elements" in the repertory grid elicitation framework) that they had experienced in the iProject, along with the characteristics of those information sharing mechanisms. The basis for their answer was bilateral: either as someone who shared with others, or as the recipient/receiver of that sharing action, or both. (One participant was reluctant to conform to the structure of the protocol, so our discussion was shorter, general, and semi-structured.)

The participants identified 295 information sharing mechanisms, with the number per interview ranging from a low of six (not counting the one that was zero) to a high of 24. Because each participant used their own words, it was necessary, in essence, to create a categorical taxonomy of “things that were shared” in order to group similar items and present a consolidated set of results. Figure 4-1 presents the top 11 information sharing mechanisms across all participants, with a small inset of the Appendix D graph of all 295 items. The frequency distribution of the information sharing mechanisms by categories (Appendix D) shows some interesting patterns of information sharing mechanisms; not surprisingly, meetings was the most frequent mechanism across all groups. Code (Software) was the second highest item, a bit surprising in one sense because less than half of the participants were performed a software development role. On the other hand, from a discipline perspective, 19 of the 23 participants have either Computer Science or Computer Engineering in the educational background, so perhaps code is a meaningful expression regardless of the project role. Also surprising is that email is in the top 11 but somewhat lower in rank than project management, design, discussion, and presentation/charts/slides. Moreover, despite its ambiguity, the two forms of demo (thing – the standalone form run in self-service mode) and demo (event – a narrated session with a story line and a visual representation of the concepts) are both in the top 11.

Not surprisingly, there were many common information sharing mechanisms named by the participants (both as “events” and as “artifacts”, listed in Table 4-1 and Table 4-2 respectively), such as “meetings”, “email”, and “demo”, as well as ones mentioned by only one participant, such as “workshop” and “visit to lab”. Nevertheless, the subjects in many cases experienced the same named element quite differently, often across disciplinary, role, location, or organizational boundaries. Section 4.4.2 contains a full discussion of this finding.

Throughout the lifecycle of iProject, but especially at the beginning, there was a dependence upon semi-structured, descriptive materials for information sharing. Taking multiple forms, such as a PowerPoint presentation or diagram, the participants described how often they were created by a single person in advance, but then used together in a group setting. This shifting back and forth between individual activity and interaction, between solitary efforts and collective activity, and between solo work and collective activity was a pattern throughout the research data. Some activities and tasks were solitary in nature. An example of this was a PowerPoint

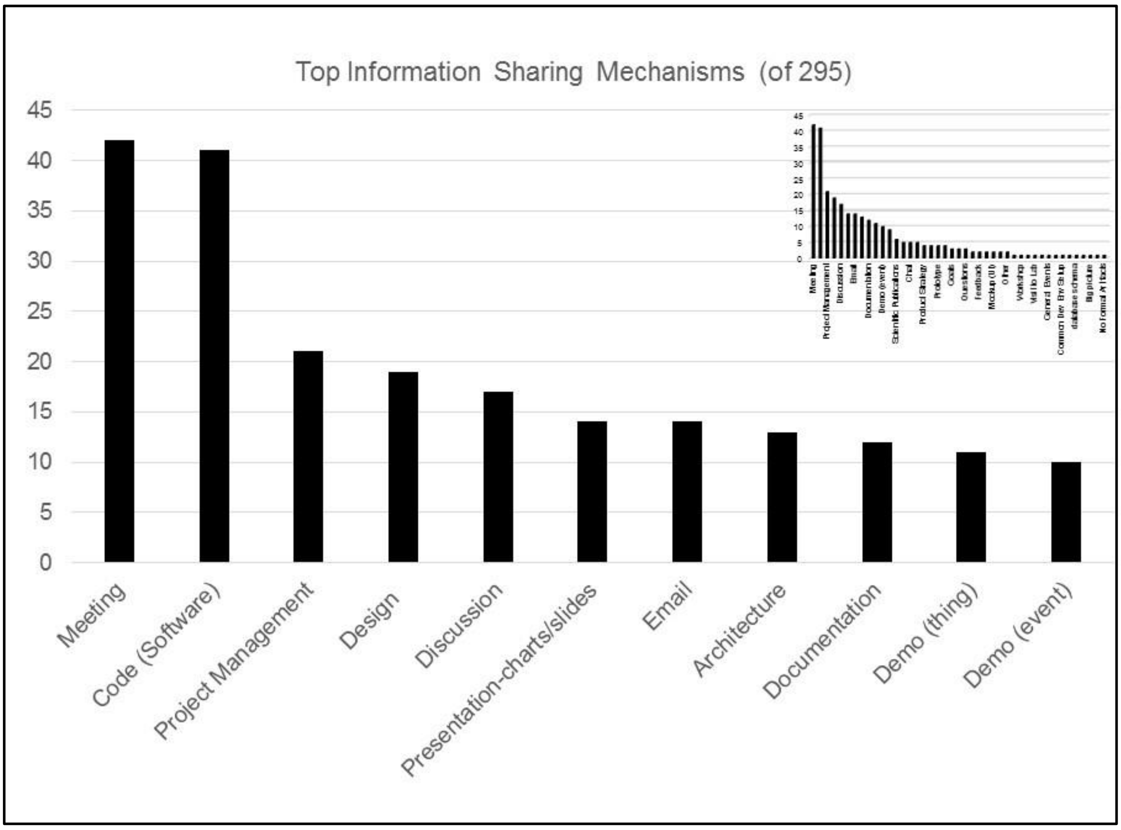


Figure 4-1: Top 11 information sharing mechanisms

presentation constructed by one person in advance, but used interactively and collectively in the discussion of a meeting. A second example of individual creation and collective use was the creation of a "wireframe" -- a type of computer interface design -- showing the user interface design and flow. Activities such as discussing one of these individually created artifacts, arguing, or negotiating and agreeing on architecture are examples of collaborative activities:

> The development of system architecture is a lot more collaborative than the development of code and systems.... I mean the collaboration is a lot more intense when doing systems architecture … it's not a solitary thing…. because at the end of it you need to come to an agreement of what the system architecture should be in terms of the code. Once you're assigned part of this, like a feature or something … it's solitary activity. (Researcher/Software Developer, RSD2, 12-17-13)

Information Sharing Mechanism (Event)	Additional description in detail
Meeting	Many configurations - all in-person, all telephone, hybrid, attending via telepresence robot. Some augmented by screen-sharing, distributed charts.
Presentation	Many configurations - all in-person, all telephone, hybrid, attendance via telepresence robot. Some augmented by screen-sharing, distributed charts.
Demo	A live event to demonstrate an idea – from conceptual to extremely specific - through showing a mock-up, prototype, or actual running system.
Discussion	Many configurations, from 2 participants to many - all in-person, all via telephone (remote), hybrid (a mix), attending via robot. Some augmented by screen-sharing, distributed charts.
Screen sharing	Screen sharing via multiple technologies, in-person, remote, and hybrid.
Workshop	A working meeting, in this case, in-person although other configurations are possible.
Instant messaging	Internal company instant messaging capability for 2 or more people. Transcripts were saved for later usage.
Telephone discussion	A verbal discussion between two or more people; may be pre-scheduled or spontaneous.
Visit to lab	An in-person visit to a workgroup or location that is not one's home working environment.
Serendipitous moments	Unplanned encounters, or unplanned moments that occur in a planned or formal event.

Table 4-1: Information sharing mechanisms: events

Information Sharing Mechanism (Artifact)	Additional description in detail
Charts, slides	Sometimes called a presentation but differentiated from event; often PowerPoint
Email	Questions and answers exchanged, directive to do something, many other usage types
Word document	
Product strategy document	Product strategy, product roadmap, customer needs
Research strategy document	Research strategy, current work, roadmap
Goal	Wide variance
Objective	Wide variance
Design artifacts	Scenario, use case (story), common use case, requirement, assumption, design idea, design hill thinking playback, joint specification
Architecture diagram	Architecture diagram, system architecture, system dependencies, system implications), Marketecture (or Marchitecture) diagram, Research solution diagram, whiteboard drawing, digital photo of whiteboard drawing

Information Sharing Mechanism (Artifact)	Additional description in detail
Feedback	Reactions, suggestions, and advice about some work aspect.
Software	Code, code extension, source code, technologies, development build, algorithm, each workgroup’s technologies, example java project, Code, Open Source Code (e.g., OpenRefine), API, user interfaces, entity graph, jars/libraries, product software, UML class diagram, running code – actual system, physical integration, brainstorming platform, defects, performance issues, meta data layer and defined extension points
Development environment	Common Development environment, CDE setup, source code repository, bug tracker
Data sets and metadata	Data set, meta data, common data set (air quality, health data), data file format, data properties, public data, raw/cleaned-up data set, database schema (in PowerPoint)
Demo	Real running system and/or play with it yourself; sometimes called a prototype.
Video	Recorded screen captures with story/narrative
Prototype, working prototype	Software implementation – expression of functionality
Mock-up	Paper, PowerPoint, and other light-weight expression of functionality
Project Management artifacts	Project plan (sometimes MS Project), milestones, status, to-do’s, workflow, sequencing, plan(s) for the future- short and long term, sprint, tasks, priorities, concerns, monthly update, next steps, support activities, scalability, staffing questions
Process	Project on boarding, research proposal, security, formal requests, dataset approval, how to manage datasets,
Documentation	Wikis, internal forum, website, installation instructions, co-authored published papers, joint patents (invention disclosure), and description from invention disclosure project repository (e.g., wiki, persistent).
People on the team	Clean boundaries between core team members, specialization, co-located, remote, teacher, core group
Expertise	E.g., UI, graph, database, WebSphere, library profile
Big picture	A high level, macro view of a system or piece of software; the purpose, overall design and architecture
Questions and answers	Questions asked of colleagues and associated answers; informal, adhoc

Information Sharing Mechanism (Artifact)	Additional description in detail
No formal artifacts	Common understanding, "gentleman's agreement". (The opposite of an artifact; mentioned as a mechanism.)

Table 4-2: Information sharing mechanisms: artifacts

There were some interesting variations between the iProject Core and Stakeholder groups in their interview responses at a high level, as shown in Appendix E and Appendix F. The two groups contained roughly the same number of people (7 in the Core group, and 6 in the Stakeholders group). The Code (Software) category had a total of 28 instances – the highest value – in the responses from the iProject Core group, who provided multiple examples of different types of Code that were shared in their project. In contrast, the highest number of instances for the Stakeholders group was Meetings at 16, reflecting enumeration of specific types of meetings for information sharing. These responses show a difference in how people looked at information sharing, favouring mechanisms that are closer to the actual work itself. For the Core Group of participants who performed the technical implementation, software development, and systems integration, code (software) was a key mechanism. On the other hand, the Stakeholders reported that the Meetings category was the highest, which again were consistent with the fact that Meeting as an information sharing mechanism is a frequent work activity of that group.

4.2.2 Information sharing: affordances/deficiencies

The participants critically evaluated the affordances and deficiencies of the information sharing mechanisms – the artifacts and events of information sharing – when sharing choices were made. One participants noted it as "something is missing", focusing on the deficiency. Either way, it was clear that the mediation of whatever tool/artifact, in activity theory nomenclature, had an effect on the sharing. There was a preference for simpler and familiar collaboration tools throughout the project lifecycle for sharing information. This is because some of the technologies were missing functions, or did not reliably provide those functions, to support collaboration in a stable and predictable manner. Participants viewed information sharing mechanisms as problematic that interrupted the "flow" of the main line of the activity, in the words of one participant:

> [The presentation is] generally pre-baked because… in-the-moment sharing tools aren't that good. They're just clumsy to use, I think, slow and awkward. I don't use any. (Researcher/Co-lead, RSD1, 12-18-13)

Participants frequently mentioned screen sharing, instant messaging and wikis as effective information sharing mechanisms. They often shared prior artifacts close or central to the task, and used them as templates for a new task, including computer code by the software engineers and presentations/charts by the architects. This was a "by example", direct usage of project artifacts; a technique of showing and doing more easily; and a technique of facilitating understanding.

There was also a preference for real-time processes. This included verbal discussions between varying numbers of people (from two people to many), real-time decision-making in meetings, and instant messaging if all parties were working at the same time. Temporal aspects were an important factor in information sharing mechanisms. Documentation often lagged, requiring the use of people as sources, or utilizing direct sources (the computer code).

> But as the implementation evolved the foundational architecture became out of date and stale. And so furthermore I wouldn't be surprised that if it's stale again because we update it as we see a need. (Researcher/Software Developer, RSD5, 7/11/14)

PowerPoint charts and live presentation delivery were often used in combination to facilitate interaction, looking forward in time; and publications and patents to publicize, and to document what was accomplished, looking backward in time.

Surprise and serendipity. Participants mentioned the unplanned both as a positive force and as something disruptive. An example of this was a discussion tangent in a meeting, seen as an interesting way to explore new ideas not on the meeting agenda.

Ambiguity tolerance. Participants noted some important terms as ambiguous in their usage. These include demo, prototype, and proof-of-concept – all examples of "showing something" or "trying something out". Often in the Demo event, people were not sure exactly what had been actually implemented, what it was made of, and what was just a conceptual demonstration. Use Case and Scenario had similar ambiguity, as there was evidence that they were either undefined in usage, multiply defined across discipline or role, or a combination of both.

Information Sharing Mechanisms	**Advantages/Disadvantages**
Meeting vs. email	Differences cannot be resolved via email: "It's always when we get back together in person in a meeting, either regular meeting, or we just say okay let's talk about this tomorrow. We come together and we agree in a meeting face-to-face. It can never be on email." (Researcher/Software Developer, RSD2, 12-17-13)
Creating in advance vs. "on the fly"	"Unfortunately these days, we're not using anything interactive so it's pretty much pre-made PowerPoint. Yes, I mean in other situations I've taken notes, you know through a meeting interface and let everyone see that. I don't think we did that much on this project." (Chief Architect, CA, 12-6-13)
Rapid prototyping tools vs. overview technical documents	Java doc and "hello world" programming examples are effectively used to get programming going quickly, but do not provide a good overview: "Sometime you try a little bit and say, 'How come it doesn't work? I mean, how come this is not the way I expected?' because - oh - because you don't even know the big picture." (Researcher/Software Developer, RSD3, 12-19-13)
Formal Process-driven methodology vs. Exploratory approach	"… It's emergent. It's settling into a particular shape. It's something that's not accomplished - fluid I guess - and any given moment it might be a little different from what it was before, so if you keep hearing information about it that's a way to keep up to date. People usually don't sit down and send out a mailing every time there's changes, well, unless it's very important." (Researcher/Software Developer, RSD4, 7-9-14)
Sketches, mock-up, vs. working prototype, proof-of-concept	"[the] lightweight nature - low effort, low cost way to do them; I can easily iterate …[vs] it's a lot more work, it's not a low effort kind of thing and … it's much harder to change your mind." (Researcher/Chief Architect, RCA, 12-11-13)
Demo and PowerPoint charts vs. published paper, patent	Looking forward, to facilitate interaction, vs. to publicize and document what was done. (Researcher/Manager, RM3, 7-14-14)

Table 4-3: Advantages and disadvantages of selected information sharing mechanisms

Table 4-3 summarizes the affordances of various mechanisms of information sharing, as viewed by the participants, and the advantages or disadvantage of one versus another.

The nature of the information sharing mechanisms was often explicitly considered by the participants as they discussed their choices in what they used, and discussed their characteristics of utility. There were circumstances where participants felt the use of a self-contained artifact, for example an architecture document in a wiki, to be advantageous over an artifact that required someone to provide explanation in conjunction with its use. Similarly, sometimes a point-to-point mechanism was preferred over a one-to-many scope if a one-to-one personal interaction was more appropriate.

The amount of resource required to create the information sharing mechanism was also a factor, with a preference often for low-effort mechanisms versus high-effort. A Use Case, by its nature, was more about the end-user of the system or software than the technology. These information sharing mechanisms ranged from the very general, with a wide latitude in their application, to the more specific and specialized. Table 4-4 provides additional characteristics of information sharing mechanisms described by the participants.

Contrasting characteristics of information sharing mechanisms	
Synchronous	Asynchronous
Requires active conversation	Self-contained
Low-effort	High-effort
All of the people are in the loop	Point-to-point
Enables immediate response or feedback	You have more time to think about it
Focused scope	Expansive, broad scope
Work product of an individual	Work product of multiple people
About the user	About the technology
Well-defined	Exploration
Very concrete	High level; vague
Interactive	To document understanding
Awkward; disrupts flow	Effective

Table 4-4: Commonly mentioned characteristics of information sharing mechanisms

4.3 INFORMATION SHARING: THE EXPERIENCE

This category of findings delves deeper into the characteristics of the different information sharing mechanisms, and the implications of using them for workgroup activities. This is about how the participants experience different ways of sharing information in their project, both a proactive sender of information and as a recipient. Also relevant is their judgment and rationale in selecting one method of sharing over another. In addition, it is about the implications of information sharing, in general, and specifically about specific ways of sharing information.

4.3.1 Theme: the continuum of information sharing and information seeking

In this study, although the terms “information sharing” and “information seeking” were not usually used by participants as labels for the corresponding activities, they did mention occurrences of information seeking in the context of information sharing. Information sharing and information seeking was seen as complementary activities, though differing in direction and in which parties are active and passive (Figure 4-2).

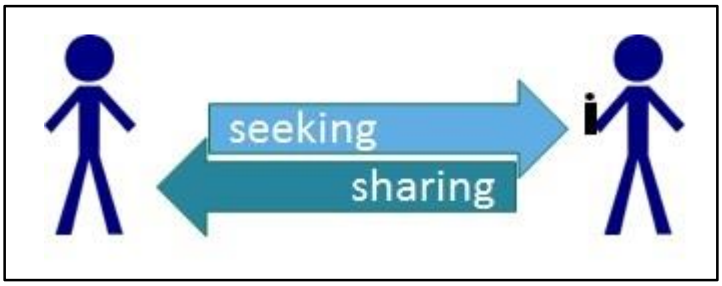


Figure 4-2: Information sharing and information seeking continuum

Information sharing and Information seeking are not opposites exactly, but activities on two ends of an information behavior (e.g., my information seeking might result in another person sharing information with me, and similarly, as I am sharing information, another person might take the opportunity to seek by asking a question). They also often occurred together in the interview responses. As one participant noted:

> … the casual conversations are… one or two people, and …we’re very interactive where I’m saying, ‘what’s the latest on...’, and ‘how’s this going?’ And … those [informal conversations] are the ones that are the most value to me honestly is to be able to really dig down and fill in the holes ...

> how is this working, and what are we really building, and what are we building it on, and so forth. (Architect, A2, 5-28-14)

An information sharing opening seemed to provide an opportunity for questions, clarification, and general information seeking – a two-way street. This was not simply a pre-planned unidirectional "transmission" of information, but an interaction that was much richer. Sometimes sharing occurred in order to open the possibility of getting something back, for example, feedback. This was an example of "giving" in order to "get".

4.3.2 Theme: thinking together with information

One participant observed that he wanted people to "think together" in his meetings. This is a powerful idea about the activity of collaboration, and one that is complex, when considering the diversity of teams and the pervasive forces at work.

The intellectual work of activities in this space occurred both in individual, solitary efforts and through collaborative efforts, with a mix of informal and formal processes. The use of information and information artifacts by individuals, in interaction, and in collaborative efforts -- tacit, embedded, explicit, shared, not shared, understood, not understood -- pervaded every dimension of this work. Moreover, it is quite interesting to reflect on the findings related to the joint intellectual efforts to accomplish the substance of the work and their enabling information events, artifacts, and instruments.

The participant who expressed the idea of thinking together talked about it in the context of describing the purpose of a particular meeting:

> **The meeting is to make sure that we are all thinking together** and cross-communicating so that everybody gets to hear the same things. During that time, people can get ideas around what else they can possibly be thinking about and driving forward because we have time to share those other ideas. (Chief Technical Officer, CTO, 7-18-14)

In addition, others expressed a similar idea:

> This is around what should be the integrated story ...we don't want .. each team to do [a] completely disjoint piece of work. So we think.. in terms of having a common story of how these different things (pieces) can come together in an integrated architecture. (Subject Matter Expert, SME, 12-13-13)

> The team[s] do not have a very close connection. We are connected by the product team. (Researcher/Manager, RM2, 5-8-14)

This brought up the dilemma of balancing the needs of an individual with the needs of the larger group, and getting a sense of how much common ground was present. As one participant remarked about a particular meeting:

> You don't want to drag 100 people through something for the benefit of one of them, but this looks like it was more of an 80/20.[80% of the people benefitted from the discussion] (Architect, A2, 5-28-14)

The working context was agile and dynamic, with emergent aspects due to many factors, including innovation and discovery. This meant that sometimes the journey and the work activities were not pre-planned. In this dynamic setting, providing the capability to participate, and capture related artifacts, was quite important. The following anecdote was shared about informal discussions and the phenomenon of “tangents” in meetings:

> If you have an informal discussion it’s informal, but it’s often basically very much around let’s clarify a concept that we think we have at the heart of this thing. Contrast that with the discussion tangent which may be driven, triggered off by something in the main line of the project, but in essence had nothing to do with it. Somebody just had an idea and they spent twenty minutes entertaining the team with it. They can be very useful, not so much to progress the established main line but … [as] part of the creative process, right? How do you basically get from: we have no idea what we’re doing to a really attractive set of offerings, and it’s not a straight line, right?
> (Researcher/Chief Architect, RCA, 12-11-13)

Another technique was the use of shared information from the internet with expert colleagues in order to fully understand and make use of it:

> So the shared knowledge that's out there on the web is very important for actually finding the solution, but once you found it, you need ways to disseminate it rapidly within your own group essentially, and have people you can go to that know about, you know, that are more expert on this thing or that thing. (Researcher/Software Developer, RSD4, 7-9-14)

There was reflection in the interviews about potential differences and nuances in the sharing of conceptual and abstract information versus the very concrete. Some people felt that very specific, contained information was easier to share over distance, or that extra work needed to be done in advance to realize more fully the ideas so that they could be grasped despite the gaps.

As noted earlier, “differing attention” of participants in a hybrid, partially distributed configuration can reduce the effectiveness of those gatherings, which is another issue. (Participants did not make any comment or observation about attention problems at meetings and calls in general.) Nevertheless, the findings here confirm that the mediation instruments introduce additional obstacles, distractions, and fidelity-reduction, and opportunities to improve the situation by adjusting components of the activity system that are problematic.

4.3.3 Theme: collaboration and artifacts

The meaning of particular information artifacts to individuals was often quite precise, but everyone, or even a majority of the participants, did not share these specific meanings. It appeared that within the shared work teams, terms referring to information artifacts were not explicitly negotiated either. For one participant, social agreement was a required factor in order to call something a “collaboration artifact”, and other artifacts were closer to the work itself, which divided artifacts into two categories. The following quotes refer to “defects” – problems – related to a piece of software or a system, and the difference between a “defect”, and a “to do”:

> Status and To-Do's are more like collaboration artifacts … [Defects] is basically an artifact of the work. There are defects all the time. They’re not work until they are interpreted and [we] say okay, this is something we need to do. That’s why it’s an artifact of what you are collaborating on [“the work”] rather than a collaboration artifact, where there’s some sort of social agreement that we need to do something about this. (Researcher/Software Developer, RSD2, 12-17-13)

Another participant expressed a similar view about the social agreement, but about a different mechanism, the Issue Tracker, a problem and request reporting mechanism used in a certain systems environment:

> Formal requests [are] represented by either the Issue Tracker, or an e-mail…. We have defined what it is that needs to get done. Or I am making the statement that something, regardless of whether it's right or wrong, that needs to get done. And I am making a very specific request, and I am putting the burden on you… it tends to be [used] more across groups, rather than in groups…. The Issue Tracker tends to be a social interface. (Technical Lead, TL, 7-11-14)

In this viewpoint, "formal requests" sent via email and the Issue Tracker were social mechanisms to assign work items to people. In the common viewpoint of these two people, information artifacts such as "Status", items in the "Issue Tracker" system, and "To-Dos" were Collaboration Artifacts because of the social agreement associated with them; "Defects" would remain an artifact of the work until promoted to a "To-Do". The Collaboration Artifacts carried a personal, and perhaps organizational, commitment for action.

Systems Architecture, APIs, and integrated codes emerged as ambiguous collaboration artifacts because of the varying social meaning and the lack of agreement about them. Comments from three different people illustrated the variance:

> Well, you look at the system implications for defining the interfaces. The interfaces also define, in some sense, allow you to draw out what the system implications are. Who is going to be responsible? Along with, at some point, they also talk about the performance that you're willing to tolerate, or not tolerate. But also, that also ends up talking about the security or insecurity of things. (Technical Lead, TL, 7-11-14)

> I consider systems architecture as the contract, and it doesn't have much of a purpose beyond that, whereas these two (code and system) are created as a result of that contract. And you run the code - I guess the system is basically the run time of the code. And it's as detailed as possible, whereas the system architecture is supposed to be higher level, abstract and not detailed. (Researcher/Software Developer, RSD2, 12-17-13)

> …this is the actual code that is being shared across locations, right. And so it's a very close level of collaboration where people in different locations are actually working on the same code base. And so in that way it involves much more coordinated effort and - to make sure that the code also works

> well with each other, that kind of thing. (Researcher/Manager, RM1, 5-7-14)

Artifacts identified as being collaboratively created included drafts of scientific/technical papers to be published, patent disclosures, presentations (first quote), and technical plans (second quote):

> So there [is] a single artifact that's collaboratively developed on some kind of rotating or merged basis. (Researcher/Software Developer, RSD4, 7-9-14)

> …doing the tasks and the plan are very fine grained. It's something that isn't necessarily produced in like a presentation by a single person or, you know, by whatever, but the consensus of the development team on what to do, and then there's an assignment … of … who does what. (Researcher/Manager, RM4, 7-11-14)

A temporal aspect of artifacts also emerged: demos and PowerPoint presentations were used to facilitate interaction in a project, with a forward-looking viewpoint. Journal articles and conference papers were used to publicize and to document what was done.

Participants also mentioned difficulties and concerns in the collaboration itself. One participant strongly stated that the use of their web services by another group was not collaboration. This circumstance was the integration of a web services interface into iProject, facilitated by sharing information via email between members from each workgroup. This denial of collaboration was surprising because the two groups successfully connected their efforts in some way for mutual benefit. It reflected a personal and individual definition of collaboration. Another concern highlighted by one of the participants was about imbalances in collaboration, whether in a hierarchical dimension or in the level of contribution. There was a sense that the situation was unequal, as in the example of authoring conference and journal papers and Intellectual Property disclosures (patents). With the works created by a cross-organizational set of people in the Research and Product groups, but the effort led and most of the writing done by the Research, some participants felt a sense of unequal contribution across the organizations. As one participant expressed:

> Well they are the co-authors and the co-inventors because ... we are discussing with them and developing the technology with

> the help of them giving the idea and bouncing the idea and those kind of things. So they are in the paper. They are in the patent. But it's like this, that when they are in the paper we will write the paper and having multiple iterations among ourselves and then sending them the sort of some kind of semi-final version to them. And they will give the comment and then that's it. So that's more of an unequal relationship that in a sense that - and because it's a paper .. [It is] in our domain. That's not [the Product group's] domain; it's a Research domain. (Research Scientist, RS, 5-6-14)

There was a curious statement from one of the participants related to collaboration. During the interview, this participant commented that they did not collaborate with the iProject team, and then declared, "I'm no longer working directly with the iProject project and never have been." While definitively negative on collaboration (along with lesser forms of cooperation), it is a forceful statement, but also paradoxical. "No longer" and "never have" do not typically go together in the same sentence when talking about the same item. How can one have stopped working with the team if they never did so in the first place? It was as if it was not politically correct to just say "I have not worked with them". This suggests that some sensitivity was touched in this interview encounter, or that the participant was uncomfortable admitting that they have not worked with this team. It is difficult to know, because the interview was startled by this statement and was unable to frame a follow-up question.

Finally, the participants often mentioned first the work artifacts as exemplars or candidates for information sharing, such as meetings for the Stakeholder group (see Appendix F), or Code (Software) for the iProject Core group (see Appendix E). The artifact that is closest to the actual work may be the best candidate to convey the idea(s), unless the recipient does not possess the domain knowledge required to utilize the mechanism, and the person sharing the information understands this.

4.3.4 Theme: 'not to share' as sharing

The decision not to share, or to share partially by holding back some information, is an important gradation factor in examining information sharing, and is interesting to explore in this study. Judgments and nuances were involved. Study participants talked about their decision not to share, or to share selectively, or some other variation, metering, or nuance. This shed light on the considerations that people thought about, such as not wanting to confuse the situation with non-essential information, or contribute to the very real issue of information overload. These judgments and decisions about information sharing had a very practical impact on collaboration and the shared work activity.

People often considered the overall objective of the work (as they saw it) when making the information sharing decisions. For example, in the preparation for the ABC demo, the Product team made a conscious and strategic decision to share less. The participant described an explicit decision to not share source code, but enable collaboration through code extensions. As the participant remarked:

> ...for this project we didn't share as much as we could have in the development environment because they actually did not have our source code when they developed which is a benefit in some ways. (Architect, A2, 5-6-14)

Less was more in this instance. Another example of the benefit of less sharing, or on the flip side, of the problems created by over-sharing, was the circumstance of receiving more than they wanted/needed and the burden/extra work that this created:

> People tend to give more information than is necessary….For example, more technical information than you need at a given time. So you have to filter out. (Researcher/Software Developer, RSD1, 12-18-13)

Information overload was a real burden here, and it may be that people had an underlying assumption that sharing is good, and that they did not consciously decide what/how to share, but erred on the side of more sharing. They may have rather unilaterally sent more content for a wider distribution to people.

Deliberate and purposeful “not sharing” by design was also described by another participant. They relayed an explicit design of work tasks to have “clean boundaries” between people, believing that more independent work roles mitigated the need for sharing:

> We have people specializing in different areas. So for example Person A is working in one area. Persons B and C are working in another area. And [we] tend to set up clean boundaries so there isn’t a lot of sharing and coordination that has to go on. (Researcher/Software Developer, RSD1, 12-18-13)

This is a form of “loosely coupled” collaboration, discussed more in Section 4.4.4 on in connection to knotworking. It is at the individual employee level between collaborators versus between collaborating workgroups. It does suggest that the cost of interdependency, including information sharing, even between co-located colleagues, is of concern and an explicit consideration in structuring work tasks.

Participants also mentioned unintended information sharing as a positive phenomenon. The following quote demonstrates the contrast between unintended sharing and purposeful sharing:

> And … there's general verbal conversations…I was thinking about things you don't set out planning to share something, but you just wind up hearing it, whereas like when we had some concern things, I might say to [my teammate] ‘I'm really concerned about how this or that is going to scale’. So there's verbal exchanges that are... purposeful. (Researcher/Software Developer, RSD4, 7-9-14)

4.4 INFORMATION SHARING: DISTANCE

Distance collaboration is an important category of research findings. The original focus for this study, reflected in the first definitions of the research questions, focused on information sharing and virtual collaboration in a software engineering context. There are indeed findings about geographical distance and related gaps in collaboration. However, as more participant interviews completed and dimensions of distance other than geography began to emerge, it became clear that the modifier “virtual” was too narrow because it addressed only one facet of distance. Multidimensional distance emerged in this complex setting, including not only geographical distance (virtual), but also ones of time zone, discipline, role, organization, and project tenure. The term “distance”, a more inclusive term for these multiple dimensions of difference/heterogeneity/diversity, frames these aspects of difference/heterogeneity/diversity in a distance context. The data from this study, from the microcosm of simple examples to higher-level narratives and a systemic view, contain evidence that multidimensional distance indeed was a factor that affects collaboration. The data analysis also suggests that there is a range of accommodation needed in information sharing to address collaboration across distance, ranging from none to mindful and effective, which is discussed next.

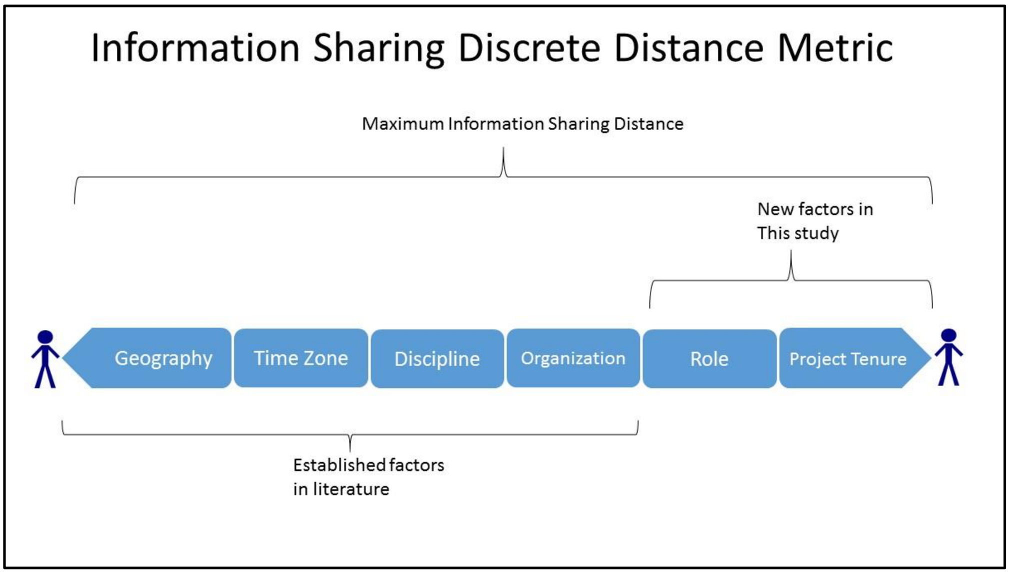


Figure 4-3: Information Sharing Discrete Distance Metric

Figure 4-3 shows a visualization of the factors of distance in information sharing, and the model of information sharing discrete distance developed in this study. The six dimensions of distance included here are geography, project tenure, time zone, role, discipline, and organization, each of which contributes some measure of distance in information sharing. For example, if two people are collocated, their geographical distance may be zero, or zero most of the time, but if they are in different countries, their distance in that dimension is greater. Similarly, if two people share a disciplinary background, they also share a foundation and ways to look at the world that have more in common than two individuals of differing disciplinary backgrounds do. The cumulative effect of multi-dimensional distance adds an increasing gap and challenge to effective information sharing.

4.4.1 Theme: context is key

As people came together to work on iProject and the related projects, each workgroup member brought his/her own context, an optic through which to understand the situation, the people, and the work. Scaffolding of this context occurred from a myriad of influences, including past formal and informal education as well as previous work experiences. However, as the project rolled forward, the context was shaped by structural factors such as the organizational structure (which may be multidimensional), how people found out about the work details and objectives of the project, and where a person fit into the structure of the project.

Every interaction related to the project also shaped context, and probably interactions not specifically related to the project too: through discussions with colleagues and management; by reading project-related information; and by engaging in the actual work, to name just a few of the important activities.

Noted here are several unexpected insights as the interviews unfolded and confirmed in the data analysis phase:

a. One participant remotely located to the core group described screen-sharing as "parachuting in… seeing the exact same thing". This was a powerful image and conveyed a powerful experience of connection to a remote person, team, or experience. Screen sharing was described as a powerful and positive mediation tool to cross physical and time zone distance, and through it to share information. This was the only tool mentioned with the capability to close the gaps of physical distance to share the same context, and to share information more naturally.
b. Although there was also "virtual collaboration" occurring with remote parties, there was an even greater amount of face-to-face collaboration among the core team members, stakeholders, and others collocated in a single location. In this localized setting, however, other "distance" issues emerged around facets other than location. These included discipline variations in formal education (e.g., Computer Science, Biology), role variety (subject matter expert, software engineer), and varied tenure on the project (original workgroup members vs. people that joined later).
c. A hybrid configuration for a meeting or discussion, with some people face-to-face around the table, and others connected via the telephone or a web conference, is a commonly used but quite troublesome configuration.
d. Also evident as the interviews progressed was the fact that this was not "a team" in the singular, stable sense, working toward a shared object and/or outcome. A dynamic topology of a networked set of collaborations and working relationships, with long-term and/or short-term shared objectives began to emerge. Each group of collaborators had one or more objectives, but they might share a common one for a period of time with another group. In activity theory parlance, it became clear that they were operating in a

"knotworked" configuration (Engeström, 1987). The Activity system constructs and analysis brought forth this insight.

Many things shaped the perspective that the people brought to their work – their education, which introduced a particular disciplinary focus along with vocabulary, definitions, processes, approaches, and methods; their previous experience; their role in the project and in their organization; and the relationship of their organization to the project, to name just some of the factors. This perspective provided a unique context in which to make sense of a situation.

Activity theory enabled capture and expression of this unique context in the activity system structure, starting from the "Subject" and pervasively all of the other points. All of these dimensions had entities that serve as mediators in the thought processes, actions, and activity: the meaning of a word, the process that comes to mind in a phrase, and the standard algorithms learned in school or on a previous project used to address a particular type of problems. The differing Subject perspectives on a particular topic in an activity system, and consideration of the other components as well, can provide insight about what is going on in a very complex system in motion.

For example, in the interviews, people expressed diverse and nuanced views of the Object of their activity (the Object in their individual activity system):

> To deliver new capabilities into our product set that can then be used to support some of the higher level department goals, as well as our customers. (Researcher/Software Developer, RSD1, 12-18-13)
>
> To help users to find the right piece of data [so] that they can do whatever analysis. (Researcher/Software Developer, RSD3, 12-19-13)
>
> We're going to make this environment, the online web environment that people who … come in ... [will] use to manage their work and get things done. (Researcher/Software Developer, RSD6, 5-21-14)
>
> I see it as the fundamental collaboration platform. (Researcher/Software Developer, RSD5, 7-11-14)
>
> Data discovery basically. (Researcher/Software Developer, RSD7, 7-18-14)

> …help people much more quickly get value out [of] analyzing data, but also helps researchers much more quickly come up with new techniques and technologies to analyze data of various kinds. (Researcher/Manager, RM3, 7-14-14)

> I see DataUI as an overarching application that allows … collaborative access. And one of the things you access is a DataLake. (Researcher/Manager, RM4, 7-10-14)

> …to encourage research, to provide researchers and research in general with the capability of being able to look at problems and do new things. And in order to do that, they require certain tools. And so we're one of the tools. (Technical Lead, TL, 7-11-14)

> …data still has to be connected, and has to be joined, it needs to be analyzed, it needs to be processed, synthesized in some way. It doesn't matter ultimately what kind of data it is. And iProject is really working on …making that process easier. (Subject Matter Expert, SME, 12-13-13)

> …a front-end to a lot of different problems one encounters with data. (Researcher/Manager, RM5, 7-9-14)

The preceding descriptions have some common components: data, data analysis, collaboration, and a generally consistent theme, but people also often described the Object closest to a facet or dimension of their own work, to their own role. For example, a software engineer described a software application; and someone working on datasets mentioned data connections and synthesis. Their focus was from their own vantage point, and the context was that their efforts were tightly linked to the larger goal of the overall project. It may be most natural for them to see their own efforts relative to the larger goal, and to shape and connect to it in this way, since they may be most familiar with their own work and understand how it connects to the bigger picture.

Another example of these differing perspectives is the polarization of responses to an inquiry about remote collaborators in the project. Some of the people in the core group had regular and ongoing contact with collaborators in locations remote to them; but others only worked with local people. It is likely that differing job roles affected the need to engage with remote people, and underscores the wide breadth in perspective: some people were able to comment on remote collaboration experiences; for others, that was irrelevant. For people working with collocated colleagues, their perspectives on both "Division of Labor" and "Community" were quite different from those involved in regular remote interactions and collaborations.

Implicit information (or tacit knowledge) held by the participants was an intrinsic factor in making sense of the shared information. Explicit discussions of the implicit foundations for ideas, concepts, or tasks were not common. An example of this is a meta-discussion about the educational background of each person, how that shapes his or her viewpoints. The definition and meaning of words and concepts is another example. People instead reported that, as extended workgroups began to work together, there was generally a gradual unfolding of a shared vocabulary and an incremental building of common ground; and usually this was an implicit process.

> So I think the way that some of these terms have had meaning is that, you know, we've traded some PowerPoint decks. We've had, you know, various face-to-face discussions, quite a few weekly calls, those sort of things where these do start to get some additional definition so really just kind of shared vocabulary at that point by working through the problem statements. [I] think the way we've typically reached a common vocabulary is just by working through different scenarios, use cases, and we kind of start to narrow down or converge on some common terminology, you know, you said this, do you mean that or, you know, those sort of things, and reach some sort of consensus. (Chief Architect, CA 12-6-13)

However, this evolutionary, emergent approach also introduced confusion. Multiple people mentioned the use of "use case" and "scenario" as important information sharing mechanisms in the design process across the organization, and used to tell a customer story. Different definitions were reportedly in play, with confusion resulting consequently. For example, the following two quotes from two different people illustrated contrasting perspectives on what a "use case" is:

> I have to be careful when I talk about scenarios and use cases … A scenario is a work flow that cuts across a whole bunch of people - all the stakeholders, whereas the use case is basically a little lower level - how a particular system is used by a single person... ...People are always confused about what the scenario is, what the use case is. [Is the definition shared?] I'm not sure. Because I've seen it used in surprising ways. (Researcher/Software Developer, RSD2, 12-17-13)

> Well, the use case is really the story. To define that, and bring in the data to support that story because you want to have a good story for the demo or for these other people when you're talking

> to Product team. It has to have a business, you know, it has to be something that they feel relevant to the business community or relevant to their different customers. (Subject Matter Expert, SME, 12-13-13)

Moreover, the organizational context of the individual shaped their perspective and context: how they see things. Looking at the task of bringing in external datasets, or open data, a researcher interested in using that data saw this as a step on the road to their research objective; an enabler to their work, which should be accomplished as soon as possible. However, another participant with a primary focus on security and processes saw this work as risk, with accompanying need for caution and risk mitigation, and careful consideration of all aspects. These two individuals were in very different places.

Activity theory enables a view from the subject's perspective, which is critical to understanding context. As each vertex (of the triangular representation of activity system) is considered from the vantage point of a particular subject, or group of subjects, it is possible to get closer to "seeing" it from that person's viewpoint, and to see the forces that are in motion at that level. It is quite difficult to see something from another's perspective, but the visualization provided by an activity system model diagram, or sequence of diagrams, made clear the variation between individuals or groups of individuals. Moreover, it enabled taking note of systemic changes over time. The activity system context and changes to the iProject activity systems over time are discussed in detail in Section 4.5.2. but deserve a small mention here.

4.4.2 Theme: manifestation of distance

Key activity theory constructs that are important for understanding the dimension of distance, while examining a full activity system, are Division of Labor, Community, and the choice of information sharing mechanisms used in mediation in the activity system. Distance emerges as a factor in multiple dimensions, ranging from geographical and time zone distance, to heterogeneity introduced in multiple discipline collaborations, to role and organizational distance arising, and to project tenure with people coming and going at different times in a project lifecycle.

Crossing of obstacles and boundaries require additional effort and/or mediation techniques. Some people worked across the world from each other and rarely, if ever, had the opportunity to meet face-to-face. Short-term changes due to travel and personal schedules also required adjustments and accommodations. For example, distance in miles (geographical) meant that collaborators will not, or rarely, meet face-to face to work together:

> This wasn’t necessarily a planned part of our development process but we actually got meet face to face at the Conference. **We had an opportunity to talk at the same time much more easily.** (Architect, A2, 5-6-14)

Time zone distance -- meaning a constrained, non-existent, or awkward overlap in work schedules that prevented or limited synchronous interaction, or necessitated technology-gated interactions -- was only another constraint. The following description conveys the experience of a half-day time difference between collaborators:

> I have a [phone] call with the iProject Core team let's say, and it's like 9:30 or ten o'clock at night [my time] and there are five people who are slowly coming into the room, and you hear them all say ‘good morning’ and they've just starting drinking their orange juice. And … there's a certain uniting factor amongst the people in the room. First of all they're in the room. They're in the same time zone. **They're sharing a lot more.** (Architect, A2, 5-6-14)

Figure 4-4 shows a concept mapping grouped by the geographic location of the participants. The colors of the large bubbles reflect temperature, as in a heat map, with red being the most strong or intense theme based on discussion topics found in the text through natural language processing algorithms, followed by orange, yellow, green, blue, indigo, and violet as the least strong theme. Theme labels identify the bubbles. The smaller grey nodes and connections (the spanning tree) represent minor concepts and their relationship to each other and the themes. The red labels represent the cohort from whom the data is associated, such as the location of the participants (remote vs. collocated).

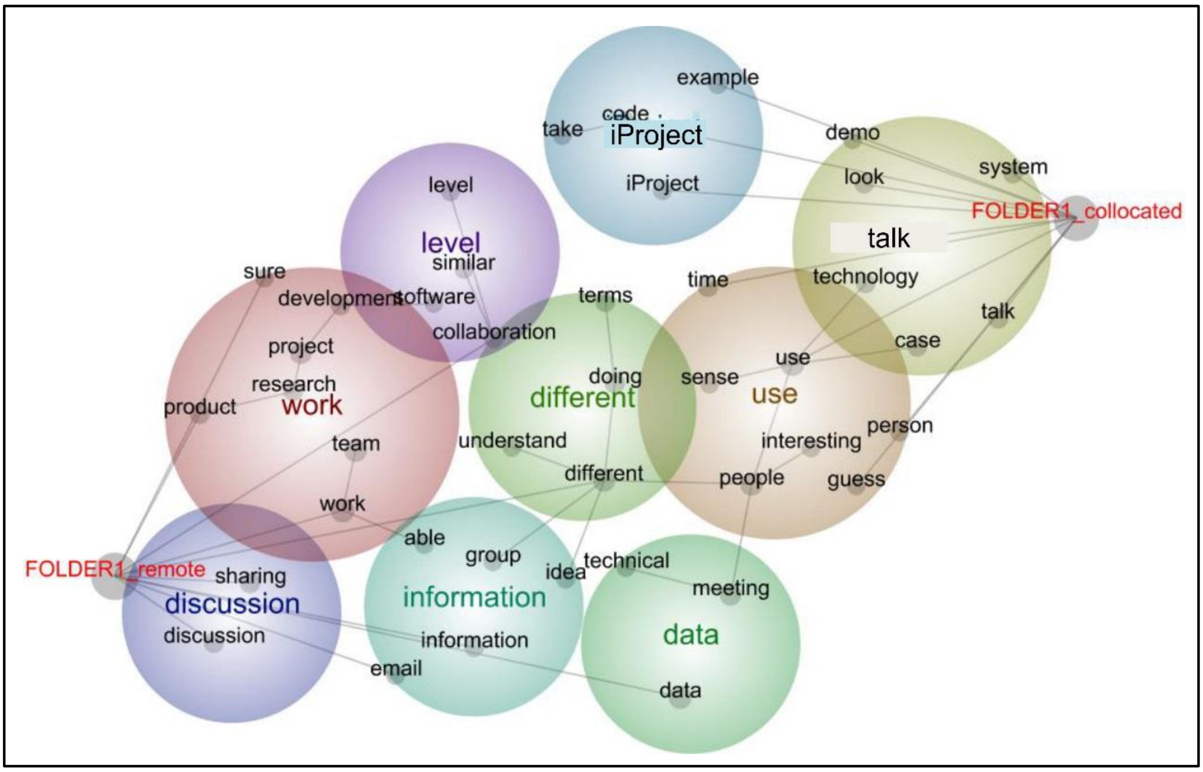


Figure 4-4: Information sharing concept map by location of participant

A verbal exchange is the closest concept for both groups. For the collocated group, it is an informal "talk"; but for the remote people, it is the slightly more formal "discussion". The collocated people also have a specific project near to them – iProject, while for the remote people there is "data", "information", and "work", an important concept for both groups.

A third type of distance was introduced by cross-disciplinary or inter-disciplinary backgrounds, meaning an educational and/or work experience context different from that of some or all other collaborators. A participant revealed the vocabulary, terminology, and language problems caused when people communicated across disciplinary lines:

> I wish that [we] had more industry people...I could relate to them a little bit better…. I **know things get lost in translation very quickly when a domain expert talks to a technical person**, and then that technical person tries to tell other technical people what they want. (Subject Matter Expert, SME, 12-13-13)

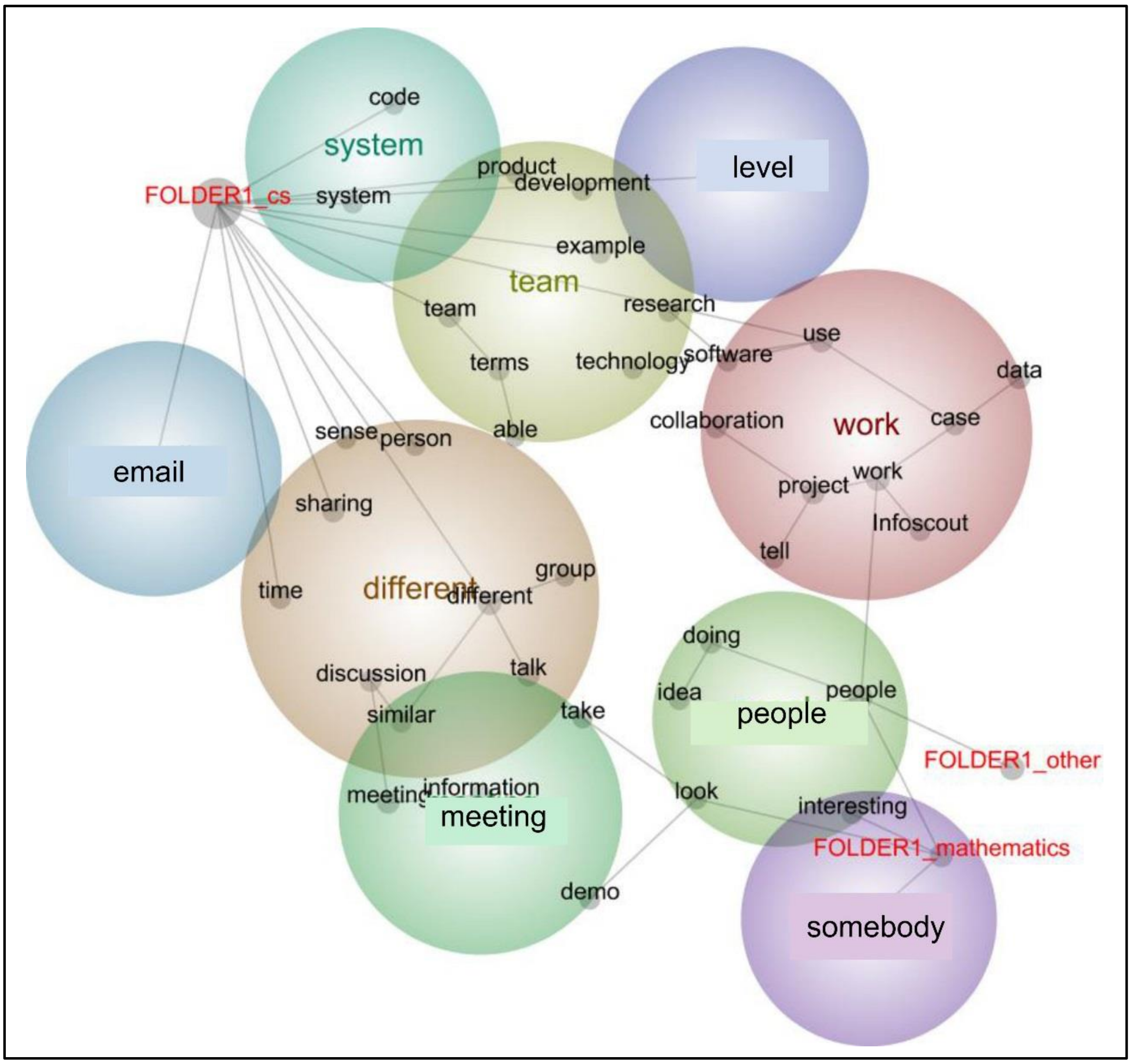


Figure 4-5: Information sharing concept map with disciplines

Figure 4-5 shows a second disciplinary difference -- differing focal attention in collaboration. As in the previous example, the red labels represent a data cohort, in this case the disciplinary backgrounds of the participants. These include computer science (noted as ‘cs’ in the figure), mathematics, and other.

This figure shows a visualization of words in interview narratives, with the discipline groupings of CS (Computer Science), Mathematics, and Other. The grouping labelled “Other” includes several disparate disciplines that were not either Computer Science or Mathematics. It shows the Mathematics and Other participants closest to the “softer” aspects: people, meetings, and the Computer Science participants deep in the technical aspects: system, code, email, development, technology, as well as team and discussion. Across all three groups, concepts associated with work were the hottest and most frequent.

Organizational distance emerged through workgroup members from multiple departments, business units, or companies, and occurred at different levels. The associated mission, goals and objectives of those organizations brought another dimension of heterogeneity to the shared activity and the shared objective among the collaborators. In their discussion, a participant highlighted how they prepared external and internal presentations differently:

> Both … [are] things that go on within the group essentially. So these are intra-group things. Whereas presentations you're usually, but not always, preparing it for the outside world for someone who hasn't been party to your everyday discussions, so you want to now put that down in a way that makes it look good and sells it to someone else. **It's more about sales in a way than it is about hardcore exchange.** (Researcher/Software Developer, RSD4, 7-9-14)

Role distance was similar (and related) to the organizational distance, but focused at the individual workgroup participants' job categorizations and their work objectives. A visualization from the concept analysis (Figure 4-6) illustrates these differences. The red labels indicate the three organizations of the participants: core (the core group of iProject researchers/developers), stakeholders, and product. Again, their individual motivation and objective may be in conflict with the project objective, or be aligned with it. Figure 4-6 shows the concepts that are closest to each grouping of participants by project job role: the core iProject members (core); the iProject stakeholders (stakeholders); the global research workgroup members from across the globe (affiliated); and the product architects from the product group (product). Not surprisingly, the product organization participants are closest to concepts groups in the theme "product"; the stakeholders closest to meeting, demo, and people, and the core workgroup closest to code, technology, and work.

Each group is nearest to components that are primary elements of their daily work. The "affiliated" participants are farther away from the technical work of iProject, with closest concepts of data, people, and example. This shows awareness and a high level of engagement, but reflects a lack of deep and intertwined "work together" with the iProject workgroup. Noticeably absent is work with code and technologies.

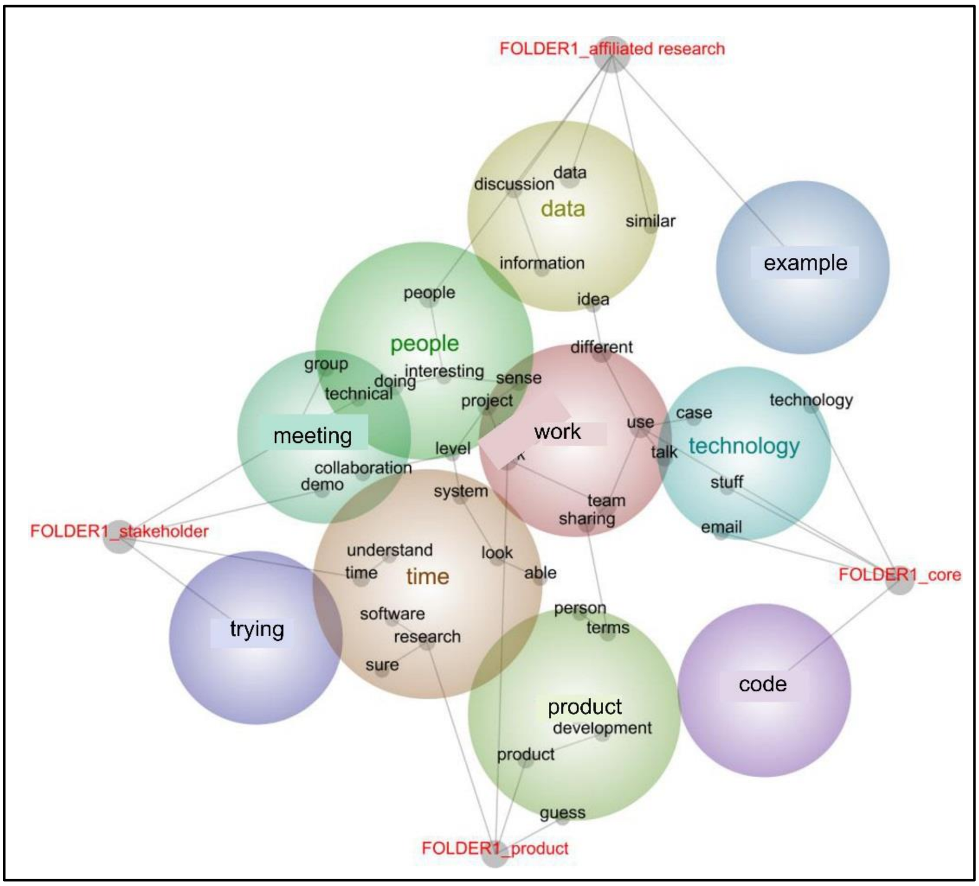


Figure 4-6: Information sharing concept map by job role

Finally, there was the distance of time: temporal distance, meaning that people joined the project at different points of time. Some workgroup members were members of iProject from the very beginning, and others joined at various points later and then needed to "catch up" on what happened before they got there. The later arrivals needed to build a good understanding of the situation at the time they joined. The agile, informal processes made this difficult:

> Right, so **I needed to understand what had been done already** and **what was the current state of it** and what they wanted to do going forward ..And so that's definitely a part of sharing of information: to get a common understanding about what is our project that we're doing. (Researcher/Software Developer, RSD6, 5-21-14)

New people had trouble figuring out what was going on when joining the project after it was underway:

> [A] photo of [a] whiteboard diagram was the first tangible evidence of structure. (Researcher/Software Developer, RSD6, 5-21-14)

One individual felt that access to more comprehensive sources of shared information, such as a wiki, would have helped speed up this project acclimation, and improved their productivity sooner. Collaborating colleagues, both within and outside of the project, were responsive in providing answers to questions and helping others to become acclimated to the project. Project members then would often produce a diagram, or a piece of technical information, to capture what they learned. Placement in a wiki or sending to others via email is the way sharing occurred.

Activity theory enables seeing the effects of distance. Division of Labor decisions in a project and the resulting impact on the other vertices in the activity system are evident in the early stages of the iProject (Figure 4-7), where one person created a mock-up and the other workgroup members provided feedback. This is a contrast to later configurations, such as the one shown in Figure 4-10, where multiple parties each had a role. In each case, they worked toward a single object, but in very different ways. In the Division of Labor decision, each activity system made an accommodation appropriate for the circumstance. In the early stage (Figure 4-7), it was a centralized, simple approach that did include collaboration, but at later points (Figures 4-8 and 4-9), it became a loosely coupled configuration that enabled widely dispersed people to collaborate.

4.4.3 Theme: emergence of new ideas and how they are embodied in information sharing mechanisms

Vocabulary was an area that seemed to be emergent within the project work, and not explicitly negotiated or defined. New terms seemed more likely to emerge incrementally and evolve in their use, rather than through a formal process or announcement. One might characterize these as “bottom up” ideas.

For example, the Data Lake concept (Dixon, 2010) is an industry term attributed to James Dixon used in conjunction with big data and data analytics activities. In the six interviews conducted in December 2013, nobody mentioned that term, and the first mention of this term was during the interview in May 2014. Eight subsequent participants in May and in July used the term “Data Lake”, indicating that the usage of the term had spread more widely in that timeframe. In early 2014, the iProject work split into two major components, with a name given to each component, respectively "DataUI" and "Data Lake", according to a workgroup co-leader. These new terms, and the exact relationship between the original project/concept and the two new concepts, emerged incrementally. There was no evidence of a formal introduction; it was evolutionary. As two of the participants noted:

> So historically there was iProject….[additional tools were being built] fast forward a few months … we all started getting very confused about what iProject was. And so, fast forward another few months, they basically came up with two more terms. To me, the seeds were there all along. (Researcher/Manager, RM3, 7-14-14)
>
> So the iProject today in my world is two things. It's the DataUI piece and it's the DataLake. (Researcher/Manager, RM4, 7-10-14)

For even a participant who could see "the seeds…all along", there was confusion along the way as a terminology evolved and was refined. For others this implicit unfolding may result in even greater ongoing confusion:

> And by the way… they claim it (DataUI) is the new name for iProject. (Technical Lead, TL, 7-11-14)

In contrast, strategic initiatives that influenced the iProject but originated outside of the project were much more likely to manifest explicitly, in the forms of email, charts, or slides, and in the forms of an announcement, accompanying information, and wider discussion about the implications.

4.4.4 Theme: Knotworking and loosely coupled collaboration

In the activity system, information instruments and the phenomenon of knotworking are of particular interest in looking at loosely coupled collaboration. People consciously adjusted both the ways they worked together and the information sharing mechanisms used to share information. Asynchronous methods were an obvious technique of choice and implemented in a wide variety of tools: email, online repositories such as wikis, documents, recorded sessions, or digital photographs. These methods enabled information sharing and movement of the discussion forward, but did not require lock-step participation by people with incompatible schedules, or differing priorities, or located over a geographical distance.

A number of distinct projects maintained their identity over the course of the study, but their workgroup members came together either as individuals or as a project team to accomplish a specific objective or outcome. Some of these “knots” were tactical, short-term, and intense in nature, and some were of a longer term in scope but periodically episodic, and casual in tone. Both sets of projects realized benefit from establishing even short-term collaborations. The demo for the ABC Conference was an example of a knotworked configuration of people who came together to achieve a successful Outcome. This planned milestone with a specific scope and due date was accomplished by a dynamic and temporary configuration of people and projects. There was a core activity of coordination and project management of the overall effort. As one of the participants pointed out, even loosely coupled efforts need a focal point.

Two approaches to loosely coupled collaboration emerged during the work to create the demo for the ABC Conference: one from the product development group enabling code extensions without source code integration, and the other from the global labs involving the transfer of intermediate, transformed data between collaborating workgroups. These intermediate results, “output” from one workgroup’s code entering as “input” into another workgroup’s code, enabled a human-aided workflow between code components. This avoided the work of actually creating automated interfaces between different code elements across the collaborating workgroups. Using different strategies, these two approaches enabled flexibility, speed, and low coordination of interdependence between workgroup members in different time zones and across many geographical miles.

> It [the extensibility paradigm] was targeted for being able to loosely develop extensions to the core platform…. it was very helpful because I was working on the extension design at the same time that they were more or less consuming it. So I could go ahead and make enhancements and changes to the extension design on the basis of these people who are already consuming it. So it worked beautifully that way. (Architect, A2, 5-6-14)

The human beings in the workgroups served as the “software” interface, by running their separate software locally and forwarding the intermediate results to the next in line:

> I didn't use a source control system because, physically, we are loosely coupled with each other. He used one file to import into the system - the file he used is intermediate result. (Global Lab Software Developer, RSD8, 5-5-14)

It is interesting to note that more tightly coupled collaboration existed in the context of an overall loosely coupled landscape as well. Within the iProject Core workgroup, there were carefully orchestrated activities requiring very detailed coordination. Although these were “informal”, the workgroup needed intertwined information flow to accomplish the task:

> We sometimes [do] ‘micro level things’ like, you know, is it okay for me to check this code in now, is it going to break your code, or are you ready to check in this thing we have to check in together. Those things tend to be very much one-on-one or informal stuff. (Researcher/Software Developer, RSD4, 7-9-14)

Another participant described a “back and forth” iterative process between three workgroup members doing interspersed individual work in various combinations. Artifacts were also created along the way, such as a pictorial representation of a graph and prototype interface code:

> So for iProject we had to build a graph of entities and how they relate to each other. So in that one, for example, [Colleague #1] talked to all of us, what are the entities, how do they relate to each other, and created an artifact which was basically this graph-like pictorial description.
>
> And then the next question was okay, let’s say we have that graph. What is the interface to it? So I … wrote the first code of the interface to the graph.
>
> And then it was back to [Colleague #1 and RSD4] where they went to implement that interface with respect to using a particular deck and using another different deck, and where they to each other talked a lot, some of this reflected back to changing the interface, right, because things are not perfect.
>
> So and then.. back to me. Now I’m not using the interfaces, but now I’m actually creating entities and relationships and querying that. So it’s all the time like - I consider development

> as like - there's like solitary time where you spend doing things, then either hand off, or more discussion, [then] iterate. (Researcher/Software Developer, RSD2, 12-17-13)

This shows that even within a tightly coupled configuration, there are both solitary effort and instances of "looser" activities occurring. This is an important factor to keep in mind while moving to the next section.

4.4.5 Accommodating distance in collaboration

Distance Type	Accommodation Strategies
Geographical	Technology: telephone conversations; conference calls; web conferences; screen sharing; instant messaging (chat); email; extension points in computer code (loosely coupled). Process: define work in greater detail; minimize dependencies; foster connection points.
Time Zone	Technology: Utilize extended hour telephone conversations; extended hour instant messaging, email. Process: define work in greater detail; adjust working hours.
Discipline	Process: Informal questions and answers; specialization.
Role	Process: Informal questions and answers.
Organization	Process: Identification of connection points, Informal questions and answers.
Project Tenure	Process and technology: Personal overview by longer-tenure workgroup member; questions/answers with longer-tenure workgroup member, use of documentation (e.g., wikis) to provide detailed project history.

Table 4-5: Distance accommodation strategies

Table 4-5 shows some accommodation strategies used to address issues caused by distance. There were some strategies and technologies such as telephone calls, conference calls, web conferences, screen sharing, and email used in multiple ways to help close the gaps. Use of "connection points" to build deeper collaborations between the groups spanning geography, time zone, and organization was another approach. A participant from the Product group described the importance of these connection points for their organization:

> [Describing quarterly research reviews] .. where this is like a presentation. So they're happening like every quarter … [with] some of our senior leaders and executives. And that is more awareness for other people. They are looking for connection points. Executives are looking for connection points because they're also trying to …get a sense of like what this [technology] can do. And they have a view that can then help

> guide us also in how can it become applicable? Because they have a bigger picture view. (Chief Technical Officer, CTO, 7-18-14)

The practical implication of these distances requires consideration in Division of Labor decisions in order to optimize the operation of the overall workgroup, and to increase the potential of project success. Structuring the collaboration as loosely coupled was an approach used in two different ways: through extension points in the code, and by manually working with intermediate data results. Another approach was to divide the work to minimize interdependencies between people. For example, a permanently remote colleague received assignments of tasks that were well defined and contained:

> We do have a once a week scrum meeting with [the remote team member] but still we do encourage them to take a generally agile-like 'see if you get this little piece working', 'You work on this now', or 'this is the most important thing to do now'. We say 'here, do the details of this and then show us what the design is, and then we say 'yes, that's good' or 'okay, tweak this'. (Researcher/Software Developer, RSD4, 7-9-14)

> Basically we needed something implemented. And we had a call with the developer saying what we needed, the timeframe, etcetera. Being able to communicate with them, understanding the challenges of communicating with somebody all of the way on the other side of the world without being in the same physical location….understanding what their … experience is with the libraries that you're using. Communicating what it is you want. And making sure that that's what they understood. (Researcher/Software Developer, RSD7, 7-18-14)

The challenges of remote collaboration were apparent: the care of maintaining regular contact (weekly), to maintain a "back and forth" exchange on the work progress, and to have effective communication. The sender needed to specify the task ([what] "we needed ... implemented") as well as to confirm the reception of the communication ("making sure ... they understood."). The additional work required in a remote collaboration was also evident here:

> You need to work a lot more. What that means is basically - so I can easily talk to you [about'] an idea and describe what that is, maybe using visuals, a [white]board, or something else, and I don't have to go through every detail. Whereas over remotely, I

> need to again do more work to lay it out more clearly. (Researcher/Software Developer, RSD2, 12-17-13)

This may have positive as well as negative implications. Maybe it was a good thing for collaborators to "think ahead" in advance and prepare their thoughts more completely rather than being purely "in the moment", or some mix in-between. Having remote collaborators may have inspired some different individual and group processes which benefitted both the remote and local workgroup, such as sending out meeting materials such as PowerPoints, diagrams, and action item lists. An individual employee may wish to optimize his/her individual workload by not having "extra" work driven by remote people, but it may be, overall, better for the project to have those remote collaborators participating in the project, to benefit from the dimensions of diversity that they bring. Moreover, the sharing of more information might bring a more immersive experience in the workgroup for all.

The implicit (or tacit) barrier of information exchange introduced by the dimension of distance (as discussed in the previous section) created some challenges to collaboration. The most basic issue is simply perception of the subtle, unspoken, partially spoken, or non-verbal information. Many technology-mediated, interactive information sharing mechanisms -- such as the telephone, web conferences, or instant messaging -- often prevent perception of those nuances. The following quote of interview narrative is an example of implicit information detected by being physically present in an environment with others:

> Sometimes you just start going around and talking to people. It's not always delivered like, you know, all wrapped up in a meeting, you just start hearing …people talking about things a little differently than they were a week ago, and then you can ask them 'hey, I thought we were doing this'. [Response:] 'Oh yes, we were, but then this happened, and now we're not'."(Researcher/Software Developer, RSD4, 7-9-14)

These subtleties and nuances of meaning can be difficult to detect during a telephone call or a telephone conference:

> There will be other humans who are present in team meetings in the room, and when you're remote it's never the same as being in the room. You know, you miss the subtleties… you'll think twice about bringing something up or, discussing something or collaborating on something. The whole process just feels a lot less natural and easy. (Architect, A2, 5-6-14)

> There have been many, many efforts made to establish something like electronic white boards and none of this works. It’s not necessarily just a technical problem - I think it’s also a social problem. It’s just not as easy to be spontaneous if you’re all sitting in different offices and looking at your computer. (Researcher/Chief Architect, RCA, 12-11-13)

In fact, depending on the exact technology used to connect a person to a remote event (e.g., telephone call or a web conference), for a remote participant, sometimes it was difficult even to hear that which was explicit.

A particularly difficult configuration for an effective conversation or meeting with more than two people was one that included at least one remote participant and two or more collocated people. This “hybrid” configuration (partially distributed teams) of some individuals co-located with colleagues but distant from others created many imbalances that negatively impacted information sharing in those settings. The collocated people had a very wide-band communication between each other: they experienced the same physical environment, could see the same thing (usually), and had the benefit of seeing faces and the non-verbal subtleties of communication. The remote people had a more narrow-band experience, as the people gathered together might not speak close enough to the microphone(s) (for the benefit of the remote people); they may have had side conversations (difficult to unravel remotely). Moreover, they may have been looking at visuals not shared with the remote people; and in fact, they may have (at least temporarily) forgotten about the fact that the remote person could not see/hear something. The following comment by a participant, a remote collaborator connected into a meeting by phone with others gathering in person illustrated their experience:

> I think it’s a general problem with projects where you have the vast majority of people collocated and a few not so... it [is] harder for people like us [remote], and we need to make more of an effort to engage in discussion. On the other side you have people in the room, and they can make an effort for the first 30 minutes always speak up and talk to the microphone, but it fades away, right. (Researcher/Chief Architect, RCA, 12-11-13)

The presentation of the details of information sharing mechanisms, as reflected in the gathered data and data analysis, is complete. Now we move on to examine the mapping of the data to the theoretical framework, activity theory, to gain insights from the application of this theory.

4.5 INFORMATION SHARING: ACTIVITY THEORY PERSPECTIVE

An activity system model is a concept and a conceptual diagram developed in the activity theory. This section describes the analysis of the data in the framework of activity theory and the accompanying activity systems.

4.5.1 Cultural-historical activity theory: sociocultural aspects of information sharing

Activity theory, in its cultural-historical tradition, provides the capability to focus on the sociocultural aspect of an activity in a context in the large, and is able to illuminate smaller aspects of a larger context, such as information sharing. The six dimensions of distance discussed as the chapter unfolds are all sociocultural aspects of information sharing in the collaborative software engineering activity. These aspects are: (1) geographic, (2) time zone, (3) organizational, (4) multi-discipline, (5) heterogeneous roles, and (6) varying project tenure. There are also additional sociocultural dimensions of the information culture that developed over time that were evident. The "Rules/Norms" is just one facet of the activity system where the sociocultural aspects unfold. They may also be present in the other dimensions where expressions exist about the history and how the past shapes the present, as well as about some related social aspects such as the community and division of labor. A discussion of several observations about this aspect continues below.

The first and most important aspect to note is the support expressed for openness in information sharing by an upper-line manager of the organization of the iProject:

> .. the example that they set in terms of information sharing by being honest and not having layers is just phenomenal, just phenomenal. (Researcher/Software Engineer, RSD6, 5-21-14)

For the study participant making this comment, the tone setting by this senior manager was very important in real terms, and in modeling support for the rest of the organization too to be open in their information sharing. This is a key component of the information culture in the organization, and it permeates all other aspects of the work activity as well.

Secondly, multiple participants mentioned the concept of "social agreement" when discussing information sharing mechanisms. Certain information sharing mechanisms (e.g., a reported software bug or an issue entered into the issue tracker) carried the weight of social agreement in the organization. This implied acceptance and responsibility to act upon these items:

> That's why it's an artifact of what you are collaborating on, rather than a collaboration artifact where there's some sort of like social agreement that we need to do something about this. (Researcher/Software Engineer, RSD2, 12-17-13)

Third is an anecdote about how casual social interactions in the larger work environment provide the openings for information sharing:

> ….social, so actually I mean talking to people in the lab is why they knew I had the file, so. Right? I mean, how do you get invited to the party, well, they know you have something that they want, so there you go. So I think people just kind of know what you are working on or what you've done, because of that social side. And so when the opportunity comes up I think they're going to think of you for what you may have to contribute. (Subject Matter Expert, SME, 12-13-13)

The general interactions between a wide collection of people – not only those in the core work activity, but also people on the periphery, or remotely located – are foundational in creating connection points on the social side that can be utilized to exchange information across a wider community.

A fourth aspect is about the dual role that many of the researcher/software engineers played. One side is the accomplishment of innovations in research work by building software – prototypes, proof-of-concepts, demos – that enable the research team to experience and understand unexplored territory. The other side is that researchers and customers will be using the software, which needs to work, although it is not a product or a production system. These two dimensions can be contradictory forces in the work activity and the culture, and a researcher/software engineer may have difficulty negotiating them. As one participant remarked:

> And so now, going forward, what are we going to do because I have to fit in. I have to fit in with the existing system. But I'm a researcher. It's in my nature to do something different than what's been done.… How do I fit in as a professional if this job is to, like, not fit in, you know. So, but these are the projects where they say, you know, the project is specified as: figure out what needs to be done. Whereas this is: do what, we know what needs to be done, do that. And that's a huge, you know, now that I think about it, that's like the main issue that I deal with every day. (Researcher/Software Engineer, RSD6, 5-21-14)

Finally, there is the 100-plus year history of the "Think" motto at the corporate level, and the company has gone as far as to brand its laptop computer as "ThinkPad". The ubiquitous "Think" signs carry symbolic meaning as both a historical company artifact, and a suggestion about working thoughtfully. One of the participants mentioned the "Think" concept in his discussion about the purpose of a collaborative meeting and the shared information activity:

> [Describing a status call (all teams) about customer situations, product direction thoughts, architecture team thoughts]: Yes, it's not structured, necessarily, I don't put an agenda out or any of that but there's a rough model to it….**The meeting is to make sure that we are all thinking together.** And cross-communicating so that everybody gets to hear, you know, the same things. During that time, people can get ideas around like what else they can possibly be thinking about and driving forward because we have time to share those other ideas.
> (Chief Technical Officer, CTO, 7-18-14)

"Think" is a deeply embedded cultural concept in this work environment, reminding people in their work activity to take the time to reflect. Moreover, in today's collaborative work environments, having the capability and the opportunity to think together is important. In order for a thinking activity to be collaborative, the internal thought(s) need some sort of external expression through verbal language, or written expression in words or diagrams, a shared expression. Thus, information sharing is foundational to the activity of thinking together.

4.5.2 Activity system transformations: development, innovation, and learning

The researcher initially understood that the iProject project was “a project” comprised of roughly 20 people with an objective to develop new capabilities for data analytics, even at the point of completing data gathering. However, later analysis of the data proved the initial understanding wrong. Instead, what emerged from the analysis was a dynamic, agile, and knotworked structure of multiple data analytics projects, with the following characteristics:

- Being loosely coupled;
- Having shared, as well as independent, Objects and Outcomes;
- Showing a sequence of dynamic changes throughout the study period, such as movement of staff members on and off the projects, re-organizing, and shifting outcome targets; and
- With periods of both intense collaboration and no interaction.

Table 4-6 summarizes the state of the overall activity system at four crucial project junctures and enables a view of the changes in each activity system component. The rows contain the state of each activity system component during four points in time: (a) January, 2013, (b) May through October 2013, (c) February, 2014, and (d) May through July 2014. Development and changes evolve looking horizontally from left to right. This visualization provides an organized view of all activity system components and the ability to hone in on key dimensions while looking at the whole. It provides multiple “snapshots” over time showing the evolution of circumstances in the project setting as the teams change, evolving milestones and outcomes, and new objects coming into view.

Component	(a) Status (Jan-Feb, 2013)	(b) Status (May-Oct, 2013)	(c) Status (Feb, 2014)	(d) Status (May-July, 2014)
Tools/Artifacts	Meetings, email, mock-up, slides, presentation.	Meetings, email, mock-up, slides, presentation, prototype code (real running system), milestone, poster	Meetings, email, prototype code, shared code repository	
Subject (Number of people and project roles)	iProject core developers (matrixed)	iProject core developers (matrixed), Subject Matter Expert, Product Architects, Affiliates	iProject core developers (consolidated), Product Architects, Affiliate	

Component	(a) Status (Jan-Feb, 2013)	(b) Status (May-Oct, 2013)	(c) Status (Feb, 2014)	(d) Status (May-July, 2014)
Object	Demonstrate project concepts and build support for project, future funding	Demonstrate project concepts at a conference about data analytics tooling, and obtain customer feedback. Continue to build support for project, future funding.	Shared: To deliver new data analytics capabilities into our product set, that can then be used to support some of the higher level goals for our department, and the new partnership mission, as well as our external customers.	
Rules	Conform to time slot, give positive impression.	Meet deadline, high quality customer interactions at event.	Balance commitments across missions	Secure internal income stream; balance commitments
Community	Analytics community, Research, Researchers/SW Developers, Stakeholders	Analytics community, Research, Product Division, external customers	Analytics community, Research, Product Division	
Division of Labor	Specific team member created mock-up, others participated in design discussions, volunteers	Core team implemented different components; Global Labs ran their own code on common dataset and contributed intermediate files; Product group created extension paradigm.	More core workgroup members added in reorg; ongoing collaboration with product group and affiliates	Continuation of core workgroup efforts and ongoing collaboration with product group and affiliates
Outcome	Scripted demo of mock-up presented at closed external meeting (invitees only)	Demo event at ABC Conference. Validated value for product group and customers	Interim milestones for LabUI and DataLake	LabUI, DataLake, and iProject functionality for old and new missions
Notes	iProject at the conceptual stage; project is emergent; very fluid.	Members of extended team meet face-to-face.	Re-organized in February 2014; Core team consolidated to direct reporting structure. DataLake concept starting to emerge.	New mission added – additional targets for work

Table 4-6: iProject workgroup activity system components:
January 2013 to July 2014

In addition to the tabular summary, Figures 4-7 through 4-10 show visualizations of the evolution of the iProject activity system and its key components over the period from January 2013 through July 2014, along with knotworked projects. Together, they provide a foundation for more detailed discussion of how the activity system transformed over that period, and how the forces in play shaped those changes.

There were numerous changes in multiple categories over this period, but some key developments are important to note. As the project progresses, the tools and artifacts utilized evolve from high-level conceptual expressions and lightweight mock-ups to running software, the formalisms of milestones and a customer event, and structured software code management in a code repository. The subjects, the people participating in the workgroup, expanded in size and increased organizational structure, as more people joined the project, and formal management of the workgroup expanded. The object and outcome of the activity also evolved, from a bounded demonstration of the concepts for a limited population to a usable prototype implementation deployed in a research setting. The division of labor and the community had some specific changes but the overall configurations experienced only minor changes. In nutshell, the iProject activity system started at an abstract conceptual level and increased in the expression and materialization of those conceptual ideas, ending up with an operational system implementation. These changes by stage are detailed below.

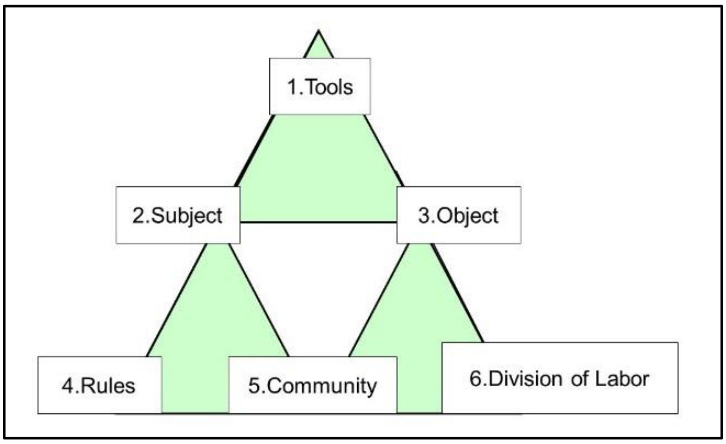


Figure 4-7: iProject activity system from January-February 2013

January – February, 2013

In this early period, the project moved from the conceptual/idea phase, with people working on it "on the side" in a volunteer mode, to a strategic, centrally managed, funded project. Figure 4-7 shows the earliest view of the iProject activity system targeting the outcome of a scripted demonstration delivered to an external audience. (This occurred prior to the start of the interviews, and the view provided by the reflections of the participants on that period.) In evidence were the foundations of the project: a strong core workgroup of researcher/developers and stakeholders, an innovative idea and interesting conceptual approach, and a clear objective.

May - October, 2013

In this next period some significant changes are already visible (Figure 4-8), although it is just a few months later. Instead of a single activity system, there is now a knotworked set of multiple collaborating projects with subjects (marked with S and representing people) in each activity system, but with a shared Outcome. We see multiple collaborations between the iProject project, Subject Matter Experts, several peer Research projects in global laboratories, and the Product group. The collaborations cross geography, time zone, organizational boundaries, and some disciplinary lines.

The October milestone was important in the life of this project because, as one of the Researcher/Developers said:

> … what it did was validate what we were doing had value for the product group and customers, and we got feedback from customers. So it was validation of an actual running demo. And the other one to compare it to would be the earlier External demo. That was a concept. People said 'oh, it's a good idea.' But it wasn't real. This is actually a real running system. (Researcher/Software Developer, RSD1, 12-18-13)

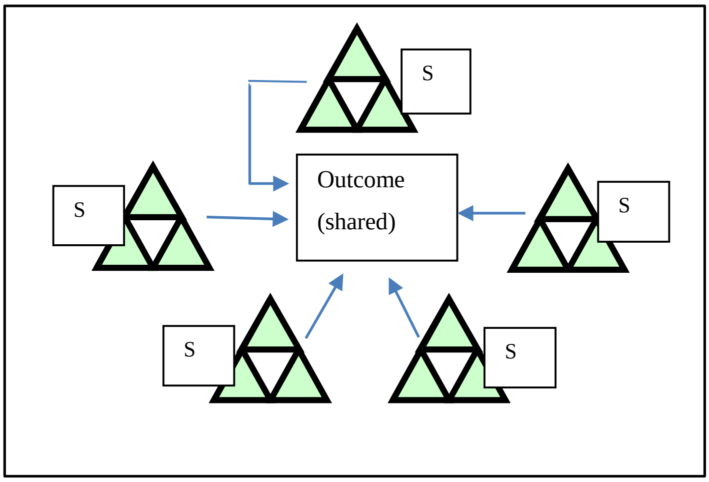


Figure 4-8: iProject knotworked activity system in May-October 2013

Development of a Use Case and selection of data to use in the demo were two strategic activities accomplished through a sequence of cross-geography telephone discussions:

> The result of that, I mean the ultimate production, was coming up with the data and use case. The use case is really the story. To define that, and bring in the data to support that story. Because you want to have a good story for the external demo or for these other people when you're talking to the Product team. It has ….be something that they feel relevant to the business community, or relevant to their different customers. (Subject Matter Expert, SME, 12-13-13)

The shared object afforded many of the collaborators the opportunity for many of the team members to meet at the conference face-to-face for the first time. This was significant in the teamwork dimension and was a very positive experience in strengthening their relationships, and having direct, face-to-face communication:

> ...We were all together in one place. And they could see and play with the demo also. You need to meet the people that you're working with eventually. So for example, the product teams we've been working with are [all over the world]. And we've gone for a year without really meeting. And so getting together at the conference was a great way to cement a relationship and have some direct communication. (Researcher/Software Developer, RSD1, 12-18-13)

Participants also reported the strengthening of trust:

> ...I mean we built up trust over the course of the year just working remotely. But actually seeing people kind of cemented it, and kind of put a stamp on it that yes, were a team. And going to move forward as one. (Researcher/Software Developer, RSD1, 12-18-13)

February, 2014

At the beginning of 2014, an organizational change consolidated the matrixed team into a single group, and named one of the technical co-leads as a manager (Figure 4-9). The very active and widespread collaboration leading up to the conference demo had ceased after the conference took place, and a new large initiative appeared. This new initiative was an additional Object/Outcome for the iProject team, and additional ongoing milestones and targets for their work. These new commitments were consistent with the general idea of the ongoing work, but brought new milestones and deadlines. The three knotworked groups in this phase shared an overlapping objective, and some overlapping partial Outcomes, but each workgroup also had disjoint objectives and Outcomes. These disjoint efforts had the potential to manifest contradictions in the associated activity system such as insufficient resources, need for shorter-term tactical solutions within an overall strategic initiative, and overlapping deadlines.

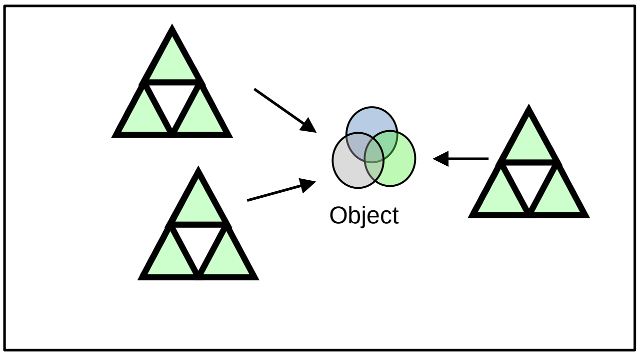


Figure 4-9: iProject knotworked activity system in February 2014

May – July, 2014

The final view (during the study) of the iProject activity system in the Spring and Summer, 2014 shows a similar configuration as the preceding period, but deepening engagement and work activity (Figure 4-10). The collaborating workgroups were the same as in the previous model diagram, but the system implementation work had deepened, and two new concepts emerged and were in usage. iProject evolved into two concepts: DataLake, a data repository, and DataUI, a user experience and interface to the DataLake.

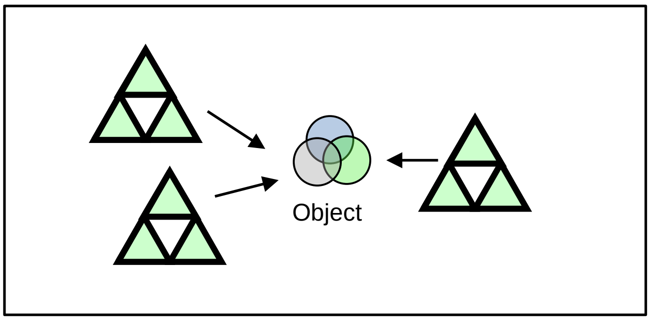


Figure 4-10: iProject knotworked activity system in May-July 2014

Multiple people also described several organizational tensions that were still at work as contradictions in the activity system, driving future changes and innovations.

These included several gaps of concern that many felt should be resolved in order to support prototype software (iProject/DataUI) for customer use. These gaps included services management disciplines, such as processes for quality assurance and testing, deployment management, and handling customer problems.

Looking at the differences between active collaborations across groups depicted in Table 4-6 and Figures 4-7 to 4-10, it is possible to see that some of the engagements across teams depicted in earlier timeframes are no longer active by the timeframe of Figure 4-10. In addition, the earlier shared Object and Outcome no longer existed by the later time.

Innovation and learning in the Zone of Proximal Development

Learning is integrated into the work experiences in a natural way, and visible in the corresponding activity systems. The results of this learning can be manifested both within the activity system of an individual person, or of a larger group. Two examples of the outcome of individual teaching and learning, one through mentoring/teaching and the other through use of a template, are included below:

> [They] sat down with me and actively wrote some code. Or sat beside me and told me which parts of the code to touch to do some initial extensions to the framework which I think is the best way to learn something. You yourself are doing it. And you get somebody to guide you. (Researcher/Software Developer, RSD7, 7-18-14)

In this example, the participant talks about an informal, interactive mentoring session where the experienced person actively performed the writing of the code with the new person, and guided the new person as they directly explored the code. In contrast, a participant conveys the following observation about a new person with model code from a colleague to jump-start his or her own task:

> AAA [a colleague] had already integrated the xyz service. So I used [their] code as a template. And then proceeded to change everything and do what was needed for the [the new] service. In the end the two codes are very different, more than alike, but I needed the template to start. (Researcher/Software Developer, RSD7, 7-18-14)

Both models are examples of information sharing, and enablement of learning using different techniques, adaptable to individual circumstances, needs, and situations. They are examples of situated learning integrated with information sharing, and with the performing of actual work. Moreover, both could help introduce innovations or new approaches as well as bring a new person up to speed in the project.

Activity system contradictions and the Zone of Proximal Development

In activity theory, a contradiction can be a positive and generative force to evolve and change one or more of the components in the activity system, manifesting as learning, innovations, or improvements in the activities and outcomes. Activity systems can have four levels of contradiction (Foot, 2014), starting with the basic contradiction of the activity system, a fundamental contradiction called a primary contradiction about the economic difference between the use value and exchange value of each activity system component (at each vertex) (p. 339). Secondary contradictions can occur between components of a single activity system, and tertiary contradictions occur when a “more advanced object replaces the current one” (p. 340). The fourth level of contradiction comes with the introduction of new practices:

> Triggered by a ripple effect from efforts to remediate a tertiary contradiction, quaternary contradictions arise between the central activity and its neighboring activity systems when a new form of practice is employed based on a reformed and/or expanded object (Foot, 2014, pp. 340-341).

There was evidence in this research that a variety of contradictions in the activity systems surfaced first as workgroup misalignments in some way, as tensions between one or more activity system constructs, such as Division of Labor, or Rules, or between knotworked activity systems.

Figures 4-11, 4-12, and 4-13 show three different instances of contradictions in one of the iProject activity systems. The lightning bolts illustrate the contradictions in the diagrams that may occur between any of the six activity system components. Figure 4-11 shows two different contradictions experienced by one participant, in a single person activity system representation. The introduction of a new tool creates a contradiction between the subject and tools component for the subject, requiring learning and proficiency building of the new tool. The second example, conformity in software development vs. the innovation of research activity illustrates the contradiction between the subject and rules components in the activity system.

Two different rules create this contradiction, the first at the project level and emphasizing conformance with the architecture and design in order to benefit from the work already completed on this foundation. The second rule, relating to the employee performance assessment criteria, emphasizes the creation of new ideas and innovations. Each of these individual examples requires growth and change by the subject to resolve these contradictions, and sometimes drives resolution to a higher level, such as the project level.

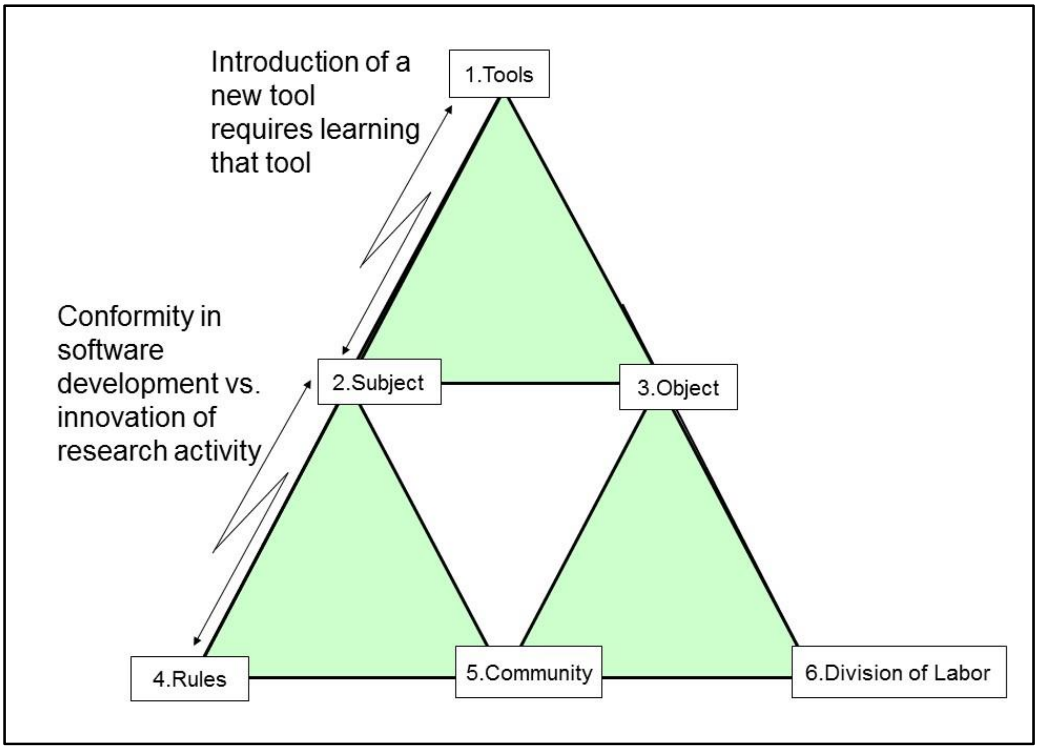


Figure 4-11: Contradictions in a single activity system

Figure 4-12 shows a wider field of view for the example in Figure 4-11. It shows the relationship between two activity systems, the original one from Figure 4-11 on the left. The object of the second interacting activity system on the right is the creation of a new strategic tool. It influences the original activity system because the object of the original activity system is to deploy strategic tools. The introduction of a new strategic tool triggers adoption, which in turn requires the subject in the first activity system to learn the new tool.

Finally, Figure 4-13 shows the activity systems for iProject and another peer project at the assessment point of determining the similarity and difference between the projects.

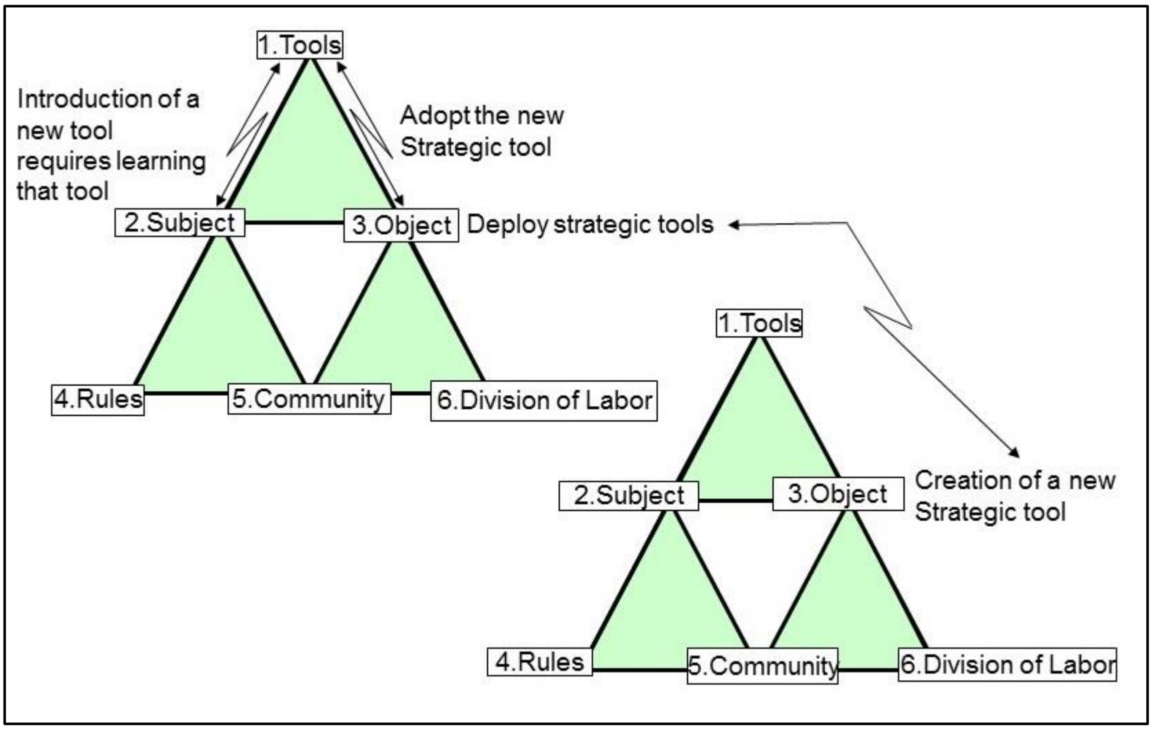


Figure 4-12: Contradictions between two activity systems

As one participant noted:

> And they have very similar claims, and describe their set of features very similarly. So there is a conversation to try and dig a little deeper and figure out what's similar and what's different, and where there might be synergies. So it's a combination of showing demos and looking at architecture diagrams. ... To figure out whether we can collaborate and have complementary function. Or whether there's complete overlap and one or the other should stop. (Researcher/Co-lead, RSD1, 12-18-13)

In the period of study, the overall activity system of the iProject project experienced at least two expansive cycles. The first occurred during the early 2014 re-organization, in the consolidation of the matrixed workgroup under one manager and addition of new people, and the second with the assignment of new mission from the peer lab. There were expansive cycles yet to play out fully after the conclusions of this study. At the highest-level activity system, these expansive cycles related to the use of the software by the customers, and the ongoing investment in the developed software making changes increasingly more expensive (in people time and other resource costs).

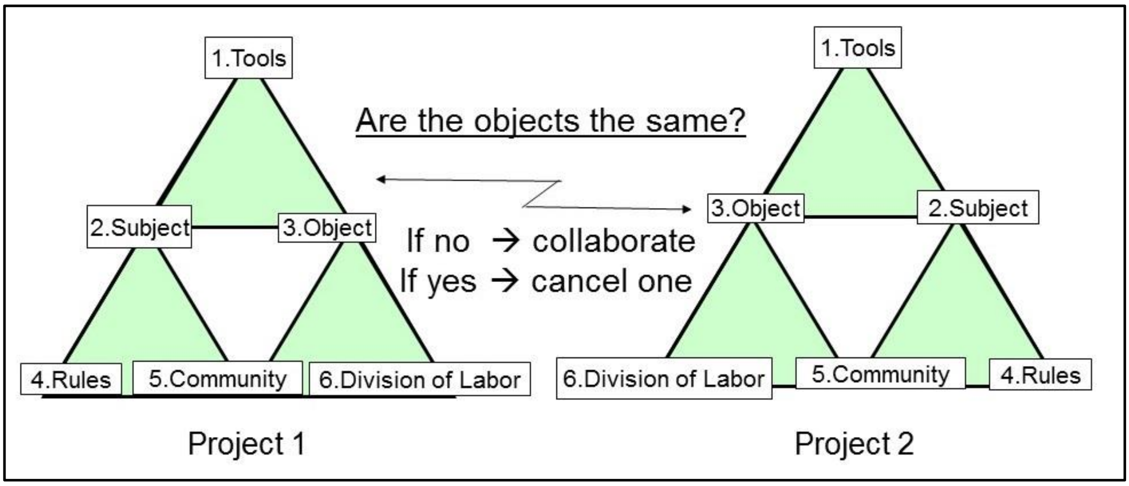


Figure 4-13: Assessment of objects of two activity systems

Activity system contradictions in smaller scopes included the change of actors from a volunteer role (relative to this project) to an in-line, core team member reporting to a manager with a consolidated team, and the adjustments of reorganization. Opportunities for future expansive cycle and transformation include changes driven by the tensions of multiple roles for many of the Actor/Subjects, such as Researcher/Software Developer, and the job role contradiction of being an independent researcher vs. fitting in as part of a larger system mentioned earlier:

> But I'm a researcher. It's in my nature to do something different than what's been done. How do I fit in as a professional if this job is to, like, not fit in. [There are] projects… specified as: figure out what needs to be done. Whereas this is: do what, we know what needs to be done, do that. … Now that I think about it, that's like the main issue that I deal with every day. (Researcher/Software Developer, RSD6, 5-21-14)

These are contradictions at the level of an individual employee, as well as systemic contradictions across the organization at the project and group level. The resolutions of these contradictions are also opportunities for growth and innovation at the individual and organizational level.

4.5.3 Operations, actions, and activities

There are different levels of human "doing" that are reflected in the hierarchical structure of operation, action, and activity in activity theory. Operations are routine, automatic, and nearly unconscious, and combinations of them comprise an action. Conversely, an action can be decomposed in order to understand it as a sequence of operations. Actions aggregate to materialize a particular activity. The examples in Table 4-6 demonstrate how work and information sharing mechanisms build from the lowest level (operation) to action and activity, moving toward achieving the objective of the activity. Specific instances in each of these three categories evolved over the lifespan of the iProject, and even prior to its start, as people brought implicit knowledge from previous experiences into the iProject. The results of those changes, reflected individually and collectively, can be examined in the development, innovation, and learning in the activity systems.

Operation	**Action**	**Activity**
The many steps in using a computer and Microsoft Powerpoint software to create a presentation slide (e.g., use of a mouse and keyboard)	Creation of a specific Microsoft Powerpoint presentation for an upcoming meeting.	Discussion of architectural options for the system under development.
The many steps in using Eclipse to create computer code and check it into the common code repository.	Creation of a specific function for the system.	Development of a software system.

Table 4-7: Operation, action, and activity

4.6 THE RESEARCH QUESTIONS

The findings have been discussed so far from several vantage points in this chapter: from the viewpoint of information sharing itself, through the activity theory and activity system lenses, from a concept analysis viewpoint, and from the angle of distance collaboration. This section summarizes the findings as answers to the research questions stated in Section 3.3.

4.6.1 Core research question: how does information sharing occur in the distance collaboration of virtual teams?

This study found that information sharing is a critical activity in the life of a virtual team, but it is multifaceted in its manifestation, and is contextual. Especially in the knowledge work of software engineering, where architecture and design concepts are often expressed first as concepts, the ability to share, interact with, and understand information is necessary and important. Surprisingly, although the initial focus of the research was on virtual teams and geographical distance, other forms of distance emerged from the study as important, too. A summary discussion of the details and characteristics of the learning from this study continues in the next sections through the five subordinate questions.

4.6.2 How do information sharing activities manifest themselves in distance collaboration?

Not surprisingly, information sharing activities permeated the process of distance collaboration, and intertwined with the actual performing of the collaborative work. Information sharing activities unfolded in both formal processes and informal actions. They were both spontaneous and planned, and occurred proactively or in response to a request. They occurred in real-time events and were resident in persistent artifacts. Every expression of information sharing had mediation to some extent – the act of taking a cognitive, mental process and externalizing it in words, body language, diagrams, and pictures. However, technology-enabled information sharing mechanisms had an even stronger degree of mediation. These included the telephone and computers, in general, and specifically via frequently used tooling such as conference calls, screen sharing, instant messaging, and email. The degree of mediation in these information sharing activities ranged from a little to a lot, and from being nearly invisible to the participants to a true barrier to understanding. They were direct products of the work itself (e.g., software code), or something that was derivative and/or created expressly for sharing. Repurposing of an existing artifact (something previously produced for another purpose) often occurred with little or no modification. Sometimes this created challenges in understanding, due to implicit assumptions and tacit knowledge that may not have fit the re-purposed use.

Mechanisms of sharing that were easiest to do, and the most efficient, were preferable. Conversely, technologies that were unreliable, or had a high threshold to set up and use, were not preferred. It was not desirable for a remote collaboration tool to consume precious minutes at the beginning of a meeting. Once the meeting was underway, it was also not desirable for usability and/or reliability issues of technology to disrupt the flow of a meeting or discussion.

The dimension of distance limited the options for, and the human bandwidth of, information sharing, particularly when geographic distance prevented an in-person discussion, or time zones that restricted or prevented overlapping work schedules, or required unusual work schedule adjustments. Additional dimensions of distance included discipline, role, organization, and project tenure, and these differences created gaps in the collaboration. Participants utilized and adjusted information sharing mechanisms to address these gaps.

4.6.3 When and in what kinds of circumstances does information sharing occur in distance collaboration?

The review of three attributes in specific information sharing activities -- time, place, and manner – reveal important details about the activity. The "temporal" aspect has at least two important manifestations in the data, the first relating to the lifecycle of a project, and referring to sharing at a particular point or milestone in the project. The beginning of a project was an important time for information sharing activities, which brought the people together through events and artifacts: using telephone and web conferences, presentations, and discussions.

Early project activities often included turn-taking introductions of people and their technologies and expertise. As the project progressed, increasing depth on common activities for the outcome of the project drove information sharing for the purposes of moving specific tasks along and providing clarification, questions/answers, and additional detail. Information sharing does occur at all stages in the project lifecycle, but did seem have different emphases at different points in the project.

A second temporal dimension is the time requirement for the individuals involved in the sharing, i.e., whether there is an urgency or specific time requirement for response. When time was of the essence, collocated people met with each other face-to-face, usually informally. The geographically dispersed people turned first to instant messaging, and then email.

"Place" is an interesting attribute, because it focuses the mind on the contrasting, complementary, or co-existing spaces of physical and virtual places. The activities of information sharing through in-person discussions and meetings were qualitatively different from those conducted virtually. In addition, as discussed earlier, the hybrid configuration with people in both dimensions is challenging for all.

Finally, "manner" provides the opportunity to consider specifics about the occurrence of information sharing. There is a broad spectrum of ways that this occurs, from the most informal and unplanned one-on-one interactions to a much more formal process of documentation and creation of prescribed artifacts within a project plan, and many gradations in between.

4.6.4 What types of information sharing behaviors and forms of information can be identified?

There were two major categories of information sharing mechanisms: events, and artifacts. In addition, they often occurred together, strengthening the experience of information sharing. The sharing of artifacts, prepared in advance, often augmented meetings and other synchronous events, such as a PowerPoint presentation shared in a meeting or a web conference. There was both a commonly practiced set of information behaviors described by the participants, as well as some less common among the participants. Interactive information sharing behaviors, including leading or participating in meetings, presentations, discussions, chat, and screen-sharing were mentioned most often, followed by asynchronous communication and written expressions such as email, PowerPoint slides and chart decks. The forms of these interactive and asynchronous communications were varied and flexible, and adapted to a variety of situations. Information sharing often provided an opportunity for information seeking, a dynamic continuum of information seeking and information sharing. The decision not to share, or to limit sharing, was a deliberate judgment and decision for the situation. Information sharing mechanisms that were closest to the actual work of the people involved were also preferred; for example, software code showing a technical approach or an innovation, or a project plans for a project manager. Demos, prototypes, and mock-ups were frequently materialized expression of abstract ideas. However, across the boundaries of discipline and job role there was often ambiguity about exactly what was "there", and what it meant.

4.6.5 What attributes are related to different types and forms of information sharing in distance collaboration?

Section 4.2.2 covers the attributes, nature, and characteristics of information sharing mechanisms. Worth noting here is the individuality of some of the viewpoints, creating a wide range of perspectives on the exactly same information sharing events and artifacts, injecting very different meaning in the collaborative use of those information sharing mechanisms. Table 4-4 in that section details the most commonly mentioned attributes that were of highest interest, although different people had very different views about their materialization in any particular mechanism

4.6.6 What purposes does information sharing serve in distance collaboration?

A variety of motivations led people to sharing information, and the fit of a particular mechanism to the near-term purpose often led to selection of a specific approach and/or tool. Table 4-8 details several common motivations reported by the participants, with broad-ranging intentions including clarification, exploration, or documentation. The object or outcome of the overall activity was often at least an implicit motivation, as well as achievement of smaller steps of progress toward those larger goals. And in this, information sharing enabled information seeking, and the answering of questions of the workgroup participants by providing an opening to explore those open items.

Used to clarify or confirm
Used to explore ideas
Used to document understanding

Table 4-8: Identified purposes for information sharing

4.7 SUMMARY

This chapter presented the findings of this study, first from the vantage points of dimensions and themes, and then replayed as a narrative response to each of the stated research questions. The analysis of interview narratives revealed a wide range of information sharing mechanisms, and many common ones were identified across the interviews. Challenges due to gaps caused by distances were evident throughout the interviews, as well as issues attributable to technologies. A variety of strategies were used to address these challenges, ranging from aligning work hours to the schedule of a remote group, use of screen-sharing, and favoring loosely-coupled work structures across distances. It was possible to see the evolution of the collaboration relationships in the knotworked activity systems, which changed as the object of the activity changed.

Chapter 5: Discussion

5.1 CHAPTER OVERVIEW

Now we move to discuss the key constructs of the study: activity theory, the research questions, and the related findings. The focus here is on confirmation of previous research, then what is new, followed by what is controversial. Next covered is the phenomenon of multidimensional distance, a critical aspect of information sharing, followed by how information sharing influences the collaboration between individuals and workgroups with distance dimension(s). Next, the nature of the information sharing mechanisms is discussed, focusing on the two-way influence between information sharing and collaboration. Two abstractions developed in this study -- a model and a metric for information sharing distance, are discussed in detail. The chapter concludes with a discussion of the innovative methodology of this study (activity theory, the repertory grid interview elicitation protocol, and Leximancer 4 data analysis).

5.2 CONFIRMATION OF PREVIOUS FINDINGS AND EXTENSION OF LITERATURE

As discussed in the Literature Review chapter (Chapter 2), the cross-discipline studies on information sharing, with either a primary or a secondary focus, provide a rich group of factors and features to compare and contrast with the findings of this study. Table 5-1 provides a summary of key factors from notable empirical studies to contextualize this study. The row shading provides a grouping of factors from a single article, and the bold font indicates a finding in one of the six dimensions of distance. This summary provides a landscape to position confirmed findings, as well as spaces for new findings (Section 5.3).

This study confirms the difficulty of maintaining common ground/situation awareness in geographic distance highlighted by previous studies (Bjørn et al., 2014; Olson & Olson, 2000; Sonnenwald, 2006). Sonnenwald (2006) found the helpfulness of informal and unexpected conversations and interactions for situational awareness and for sensing incremental change, and participants in this study reported these aspects as well.

Factor/Aspects of information sharing	Previous Studies	Found pattern/relation (with distance-related effect in **boldface**)	Source of Findings
Common ground **Situation awareness**	Bjørn et al. (2014); Olson and Olson (2000) [3]; Sonnenwald (2006)	**Negatively affected by any of the 6 dimensions of distance; cumulative effect much worse**	**Analysis of data – confirmed and augmented with 6 distance dimension model.**
Coupling of work (tightly, loosely); tightly coupled favoured to help collaboration		**Range of example configurations observed in the workgroup; loose coupling preferred over distance**	**Analysis of data – alternative conclusion reached. Augmented. by consideration of purpose in selection of loose or tight coupling.**
Information sharing has to be explicit in geographic distance	Sharp et al. (2012)	**Remote people don't understand enough about what is going on in the workgroup**	**Reported by participants - confirmed**
Individual decisions regarding what/when to share		Deliberate judgment about how much to share, including none	Reported by participants - confirmed
Technology difficulties		**More pronounced difficulties with collaboration technologies with dispersed geography in play, and across time zones**	**Reported by participants - confirmed**
Information sharing openness	Mesmer-Magnus et al. (2011, 2009)	One organization noted as very open due to leadership of the leader, and open, sharing tone set	Reported by participants - confirmed
Organizational context	Wang and Noe (2010)	Organizational context	Reported by participants - confirmed
Active and passive information sharing	Robinson (2010)	**Passive sharing occurs more frequently for collocated people**	**Reported by participants - confirmed**
Human sources of information		**Described across all dimensions, but accomplished via technology for remote people**	**Reported by participants – confirmed and augmented with nuances of technology-mediated information sharing.**
Non-human sources of information		**Use of documentation, wikis.**	**Reported by participants across all dimensions - confirmed**
Problem-solving and decision-making intertwined with information searching		Problem solving and decision-making intertwined with information seeking, sense-making, and information sharing.	Analysis of data – augmented. All collaborative aspects of work activity intertwined with information sharing
Continuum of information sharing and information seeking	Poltrock et al., (2003)	Performing information sharing in order to seek for information (reciprocity, exchange)	Analysis of data - confirmed
Organizational proximity	Wilson (2010)	**Confirmed as a relevant factor to consider in information sharing activities**	**Analysis of data - confirmed**

[3] These two articles focus on collaboration, so the purpose is to validate these factors for information sharing.

Factor/Aspects of information sharing	Previous Studies	Found pattern/relation (with distance-related effect in **boldface**)	Source of Findings
Specific sharing (1-1) General sharing (1 to many)	Haeussler et al. (2014)	Examples of sharing in one-on one settings, and in meetings, web meetings.	Reported by participants – confirmed.
Asynchronous activities, Synchronous activities; Co-located collaborations, Remote collaborations; Loosely coupled activities, Tightly coupled activities; Intragroup collaboration, Intergroup collaboration; Direct collaboration, Coordinated activities	Talja and Hansen (2006, p. 122)	**Email, wiki.** **Meeting, chat.** **Core team.** **Core team with remote collaborators** **ABC Demo** **LabUI, DataLake development** **Within the iProject team** **Between iProject, Affiliates, and Product team.**	**Reported by participants – confirmed and augmented with specific contexts and activities of the identified factors.**
Recognizing differences in the underlying meanings of shared symbols. **Sharing implications of information,** Technological influences	Sonnenwald (2006, p 6-9)	**Cross-boundary differences were noted. Ambiguous usage of demo, prototype, mock-up, as well as use case and scenario.**	**Reported by participants – confirmed.**
Person sources	Talja (2002)	Expertise was sought out.	Reported by participants - confirmed
Documentary sources		Written material was sought out.	Reported by participants - confirmed
Social sharing		**Remote participant time zone disharmony; caused feeling of isolation from group.**	**Reported by participants - confirmed**
		Social interactions informed knowledge of dataset and resulted in collaboration request.	**Reported by participants - confirmed**
Face-to-face meetings to address geographic distance issues	Powell et al. (2004)	**Meeting in person at ABC conference was important for many.**	**Reported by participants - confirmed**
Documents as multidimensional objects	Pilerot (2014)	**Information forms are smaller than a document (e.g., chat).**	**Reported by participants – augmented with additional forms**

Table 5-1: Mapping of study results into information sharing factors from the literature

The findings from this study show that geographic distance is not the only cause of “common ground” difficulties. In this study, the lack of common ground surfaced as an issue in multiple additional ways, from the subject matter expert wishing for more people of the same background to talk with about the project, to the divide felt by a remote participant for whom the meeting was in the evening but morning for others. The lack of common ground was also evident in a meeting with most attending in person but a few via telephone, where attention could wander and people in the room forgot about those on the telephone. Common ground difficulties were symptomatic in the observation that new members of a project struggled to find information on what occurred before they joined, or the struggle of a non-programmer to comprehend the meaning of a demo.

This study observed the continuum of loosely to tightly coupled work (Bjørn et al., 2014). In preparing for the ABC demo conference, the widely dispersed team (across geography as well as across organizations) chose to integrate their work by running each of their code components locally and sending the results of the processing in a results file. Participants described this integration by data file content, and aided by human processing, as an excellent approach for the objective (in activity theory, the object) and one that worked very successfully.

However, there are differing views about this in the literature. Bjørn et al. (2014) assert that tight coupling is necessary in geographically dispersed work in order to force the people in the work activity to commit to working together. This study found loosely coupled configuration to be preferable in this instance for its focus on specific objectives (object(s) in activity theory) and outcomes, and for its provision of a flexible and dynamic work environment. Additional studies on this issue will likely provide new insights. The situated nature of group work and the embedding of information sharing within the context of virtual collaboration suggest that there is likely a range of different approaches to align project setup and the associated process of information sharing with the purpose of the task. Activity theory, with its structure for evaluating division of labor in the context of the overall activity, is an excellent theoretical framework for further examination.

Confirmed in this study is the need for explicit information sharing with the geographically distributed collaborators (Sharp et al., 2012). In addition, technology difficulties (Sharp et al., 2012) with collaboration tools relative to stability and capability across the work settings (office, home, etc.), as well as user experience requiring a high threshold of effort, were also reported in this study. The participants reported being most satisfied with simple, familiar tools for collaboration (chat, texting, email, screen sharing). The technical teams had adjusted their work activities to use a set of collaboration tools, but there is definitely room for improvement. (We are still a long way from having a similar experience "remotely" and “in person”.) Confirmed in this study also is the need for explicit information sharing with the geographically distributed collaborators (Sharp et al., 2012).

The study participants confirmed the influence of two organizational factors, leadership tone setting around openness in information sharing (Mesmer-Magnus et al., 2011; 2009; Wang & Noe, 2010) and organizational proximity (Wilson, 2010). They are examples of factors outside of the core six dimensions of distance, in the outer ring illustrated in Figure 5.1.

A finding about passive sharing of information, more prevalent for co-located members of the workgroup than for remote members, confirms Robinson’s (2010) study. In addition, evidence of both human and non-human sources of shared information was also important in the sharing activities of many of the study participants.

This study found the continuum of information seeking and information sharing (Poltrock et al., 2003) in evidence in this setting, and the preferred use of people as an information source (Talja, 2002; Robinson, 2010). Haeussler et al. (2014) noted a distinction between one-to-one and one-to-many sharing, and both forms of sharing emerged in this study. All ten facets of information sharing identified by Talja and Hansen (2006) in their discussion of the characteristics of information sharing -- along the variations of synchronous/asynchronous, co-located/remote, loosely coupled/tightly coupled, intragroup collaboration/intergroup collaboration, and direct collaboration vs. coordinated activities -- were reported by participants in this study, and the first four pairs of facets (in bold) showed distance effects.

Sonnenwald (2006) suggests that the sense-making and usage activities related to shared information (the implications of the shared information, and understanding the symbolic meanings) are important, which is supported by the findings of this study. Negotiating meaning for terms and the emergence of new terms were described by multiple participants, as described in Section 4.2.3. This suggests the unevenness in concept usage among the collaborators through the project’s lifespan. Furthermore, the findings from this study show that the difficulties with understanding of terms and sense-making of information shared are particularly pronounced across any dimensions of distance.

Information sharing activities across different types of channels (formal, informal, and social) and different types of information (technical and documentation), as noted in Talja (2002), were also seen in this study, with only social sharing impacted by dimensions of distance. Finally, the benefit of meeting face-to-face to enhance collaboration across geographic distance (Powell et al., 2004) was confirmed in this study. A study participant commented on the experience of meeting collaborators in person after working across geographic and time zones:

> ...I mean we built up trust over the course of the year just working remotely. But actually seeing people kind of cemented it and kind of put a stamp on it that yes, we're a team. And going to move forward as one. (Researcher/Software Engineer, RSD1, 12-18-13)

This study confirmed the findings of Powell et al. (2004) related to the benefit of remote workgroups meeting together face-to-face. Multiple participants reported how helpful it was to meet in person, at least once, as they did at the ABC conference.

Finally, Pilerot (2004) discussed the embodiment of information in document form, as expressions, or materialized in physical form. Documents are certainly very important in many circumstances and settings of information sharing, but this study found additional forms of written expression. These included smaller and more informal forms in writing (such as email, texting, or chat) as well as events which may not exist in a narrative form, and which also may or may not be saved (e.g., transcript, video recording). The findings of this study suggest that these additional forms be included in future studies of information sharing.

5.3 WHAT'S NEW IN THIS STUDY

5.3.1 Framework for Core Discrete Dimensions of Distance

The analysis of data in this study shows that when information sharing in a collaborating workgroup involves distances, particularly multidimensional distance, those distances should be understood and explicitly acknowledged. The analysis also shows that consideration of the information sharing mechanisms across distance(s) are important in ensuring better understanding of the material by the recipient.

Figure 4-3 (in Chapter 4) presents the information sharing discrete distance metric developed in this study, a model of total cumulative distance between people, and Figure 5-1 (below) shows a visualization of the factors of distance in information sharing also developed in this study. The six core factors of information sharing (geography, project tenure, time zone, role, discipline, and organization) each contribute some measure of distance in their corresponding dimensions in information sharing.

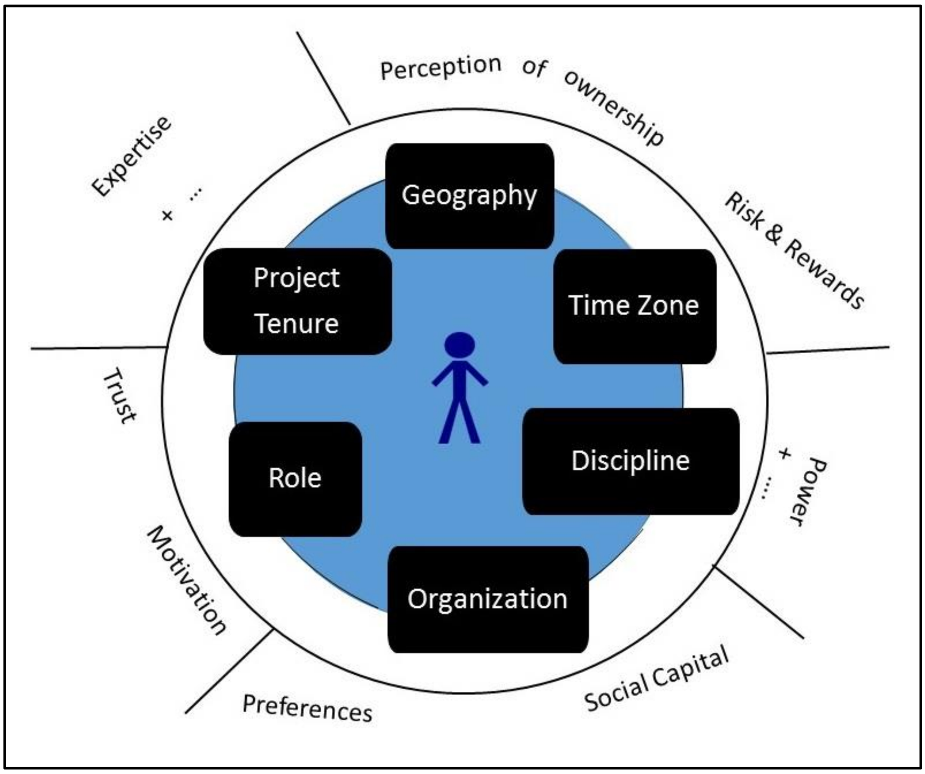


Figure 5-1: Information Sharing Distance Model

For example, if two people are collocated, their geographical distance may be zero always or most of the time, but if they are in different countries, their distance in that dimension is greater. Similarly, if two people share a disciplinary background, they also share a foundation and ways in looking at the world and experience less distance in this dimension. Consequently, they would have more in common than if they were from different disciplinary backgrounds. The cumulative effect of multidimensional distance adds an increasing barrier to effective information sharing.

Besides the six dimensions of distance included in the model, other important factors influence information sharing. These additional factors influence the experience of collaboration in the large and information sharing in particular. However, the examples identified in Figure 5-1 (e.g., Expertise, culture, trust) have the characteristics of being more abstract, more difficult to discuss, and potentially more awkward to deal with in a work setting, and the experience of these factors may be more variable in any situation even for an individual person.

5.3.2 Heterogeneous roles and project tenure in information sharing

From analysing the empirical data, this study also identified two aspects of information sharing in teamwork activities with scant focus to date: first, the effect that heterogeneous roles had on information sharing, and second, the effect that new members joining a work activity had on information sharing. This study, with the benefit of visualizing multiple contexts from the vantage point of different subjects, suggests that team members' project roles affect information sharing. Similarly, since in the real world setting people frequently join and leave projects while they are underway, the early participants in a project will have a very different breadth and depth of knowledge about what has occurred in the project compared to a late joiner. Nevertheless, further empirical investigation is necessary to gain more insights about these dimensions.

5.3.3 On activity theory

In his reflective and influential work on activity theory, Davydov (1999) summarizes eight problematic aspects of activity theory from a theoretical perspective. He makes two recommendations of relevance to this study:

1. Under the notion of "organizing interdisciplinarity" (p. 49), he notes that cross-disciplinary study utilizing activity theory is necessary and important to overcome the silos and divisions of the individual disciplines. This research also suggests using activity theory to study inter-disciplinary or cross-disciplinary phenomena.
2. He also recommends conducting longitudinal studies using activity theory in order to view the phenomenon unfolding. The movement and changes in an activity system, and the developments and transformation from contradictions, all take time to happen. The 16-month period of this study provided enough time for some of these changes to occur.

5.3.4 Reflections on 'object' in an object-oriented software world

There is a unity of "object" in the agile methodology (the actual integrated software under development) that resonates with the "object" in the activity theory framework (the purpose or goal of the activity, an objective in a general sense). The code being developed is an object in the work of the technical team, and usually an object-oriented one in today's programming languages. The interview data showed that it is a preferred shared artifact by the technical team, close to the actual work and thus conveying rich information in sharing.

When that "object" -- the code that the technical team is working on together – is considered in the context of activity, it is also an embodiment of the object in the activity system. One participant saw a difference in the nature of artifacts utilized as information sharing mechanisms in this project: a difference between an "artifact of what you are working on" and a "collaboration artifact". But there was also evidence of the software engineers/researchers utilizing code to share and convey information. This is a dichotomy, and an idea to consider more in additional studies. Is there an experiential difference in their ability to be experienced as information sharing mechanisms? Nardi (2007, as cited in Pilerot, 2014) observed that:

> …the term 'object' continues to bedevil activity theory (p.6). She poses the following question, which illustrates the ambiguity in Widén-Wulff and Davenport's (2007) article: "Is an object a motive or is it a material object toward which activity is directed? (p.2).

From this research – one that utilizes activity theory in studying an agile software development activity, the answer to the posted question seems to be 'it is both'.

5.4 UNEXPECTED FINDINGS AND IMPLICATIONS

The study also yielded some unexpected findings not anticipated at the beginning of this research. The original focus for this study was information sharing and virtual collaboration in a software engineering context. There are indeed findings in this research about virtual collaboration centered on geographical distance and related gaps in collaboration. However, early in data gathering, it became clear that the modifier "virtual" was too narrow because it tended to address only one facet of emergent distance. Visible distances between workgroup members sitting right next to each other emerged in the data, for example, different understandings of specific vocabulary between a biologist and a computer scientist in a face-to-face meeting. Multidimensional distance issues emerged in this complex setting, not just geographical distance (virtual), but also issues caused by the factors of time zone, discipline, role, organization, and project tenure distance. (The term "distance", used here as an umbrella term, is a more inclusive term for these multiple dimensions of difference, and heterogeneity, and diversity). Another unexpected finding is about innovations in the workgroup, arising from the crossing of disciplinary boundaries or other dimensions of distance – such as application of methods and models from another discipline in a new way; or the healthy discussion that take place across job role or organizational lines about goals and approaches, rooted in the different perspectives that these contexts provide. Harnessed with the right attitude and approach, these contradictions can be a force for innovation and positive change. These additional findings speak to the strength of the method utilized in the study, and are candidate areas for additional future research. These will be discussed more in Section 6.5

5.5 ACTIVITY THEORY, REPERTORY GRID INTERVIEWS, AND LEXIMANCER 4 ANALYSIS

Activity theory does not have associated methods, which creates difficulty in conducting empirical studies. In the literature, activity theory studies frequently utilize ethnographic methods. Allen et al. (2011, p. 786) notes the lack of "operationalization" of activity theory in the field of information behavior at this early point. Activity theory researchers have developed several activity theory-based techniques: an activity checklist (Kaptelinin et al., 1999); the analytical process defined by Boer et al. (2002); and Mwanza's Eight-Step-Model (2001). These three tools/techniques were useful for thinking about the study data in conjunction with the constructs of activity theory and were very helpful as guidance in this study.

However, recognizing that additional operational "tooling" is needed for activity theory studies, two tools utilized in this study are recommended. First, the repertory grid interview technique is recommended for data gathering as an alternative or to augment ethnographic methods. For the participant population of this study, repertory grid was an unfamiliar technique, with both a strength and a weakness as to the study protocol. It was a strength simply due to unfamiliarity – participants were invited to think in a new way about the subject at hand. And it was a weakness for exactly the same reason – the unfamiliarity caused confusion at the beginning of the interview, and participant questions. This process was very different from, for example, being presented with a list of information sharing mechanisms to rate on a Likert scale, or providing feedback to a specific set of questions posed by the interviewer. On balance, it was more of a strength than a weakness. Second, the Leximancer 4 conceptual and thematic software analysis tool provided automated analysis while staying within the natural language source data framework. This machine-learning tool enabled sophisticated text analytics for unstructured text, allowing the application of stop words, synonyms, word stems, adjustable granularity of text block analysis, and insightful visualizations. The pairing of these techniques provided a strong implementation of the user perspective and context – from initial data gathering all the way through data analysis – which preserved the principles of activity theory. Table 5-2 shows the aspects of the tool and method that are compatible with activity theory.

Recommended tools for use with activity theory		
Purpose	Description	Benefit
Data gathering, interview technique	Repertory grid: elicit elements and constructs	Participants name things; preserve their own language and worldview
Data analysis	Leximancer 4: semantic mapping tool utilizing natural language	Machine learning techniques preserve natural language

Table 5-2: Methods/tools naturally compatible with activity theory

The use of repertory grid and Leximancer 4 with activity theory maintains a consistent vantage point on the subject: their terminology, the meaning and concepts they describe, and their worldview. While it is true that language itself is a mediator to the thoughts and ideas, the research design of this study consistently supports the theoretical intent of activity theory in describing a context, and is recommended to others conducting similar studies.

5.6 LIMITATIONS OF THE RESEARCH

This section provides an assessment of potential weakness and limitations of this study, including careful reflection on implications for the validity and interpretation of the findings.

An important general observation is in order before getting to the specifics. This is a study of a very complex endeavor, in a real-life industrial setting, with many external and internal influences and many moving components. A variety of events and changes occurred during this period that affected the people, organizations, and technologies. The research data gathering took place over an eight month period for the 23 phased interviews (and extended to 19 months of project experience through the memory and narrative of the participants). It is, of course, not practical or even possible to consider accounting for all possible interactions, changes, factors, and influences to the people, projects, organizations, and technologies that might have some applicability to information sharing for the iProject during this timeframe. The purpose of this study is to gather data that illuminates these activities, and the experiences of the participants, in a deep, rich manner. However, these additional influences may cause false or incomplete interpretations of the data, or analysis that is incomplete due to these other factors. A discussion of these possible weaknesses in specific categories follows.

Role of the researcher

It is important to note that the researcher is a member of the larger organization studied and this is both a strength and a weakness. The researcher personally knew many, but not all, of the participants, and this may have had either positive or negative unaccounted effects. The collegial relationship might have influenced the candor of the participants in the interviews, either to improve or inhibit it. On the positive side, the researcher possesses detailed knowledge of the work activities, as well as the nuances of the work context. However, it is possible to bring in unintentional bias into the data gathering and analysis processes, and this can raise a concern about objectivity. Although the researcher made proactive efforts to maintain a neutral stance in gathering and analysing the data, unconscious subjective bias is always a concern, resulting from familiarity with the work environment and workgroup culture. To mitigate and minimize this potential problem, the researcher utilized data gathering techniques that emphasized the vocabulary and perspective of the participants (repertory grid elicitation), compared and confirmed data across interviews during data analysis, and utilized Leximancer 4, a machine learning/text mining tool that also preserves the language usage of interviewees.

Selection of the project case

Two factors were important in selecting the specific workgroup (project case) for the study. First, the technical work was an innovative and interesting software engineering project with geographically dispersed members. The second factor was a willingness from the leadership of the team to participate in the research project. The characteristics of the workgroup were a good match to the research questions and the overall research problem area. Out of scope for this study was how long the workgroup members had worked together. Another limiting factor is that the workgroup was from a single company, from two major organizations from the company physically located across global sites. It is possible that an analysis of data from a single case is applicable only to this type of settings, so the research findings should be subject to further testing in other settings.

Choice of time window for data gathering

The data collection timeframe was for a finite portion of the overall project lifecycle. The dissertation schedule provided the overall timeframe for data gathering, which was after the project had started. Moreover, this study concluded before the completion of the multiyear project. This was a pragmatic choice because it is very difficult in an industrial setting to catch a real project at the very beginning, and similarly, to be able to stay with it to the completion of the project. The practicality of being able to study a workgroup with the right characteristics, even for an eight-month period, made it a worthwhile selection. This is a limitation in that it would not be possible to see a clean beginning and a clean end of the project. However, the period was long enough to see changes emerge, and participants described events that happened earlier in the project.

Thus, as is typical in a real-world environment, the project activity was already underway when the data gathering began. The interviews were conducted beginning December 6, 2013 and were scheduled through July 18, 2014. Participants relayed what had happened from the beginning of 2013 through July 2014, a nineteen-month period. Recalling events that occurred in the past might not have had the freshness in memory than the more recent past; describing events from a year ago might not have been as precise as what happened last week. However, to address this concern, the researcher verified dates of events, and performed cross-interview analysis across the participant cohort. Moreover, it is always a possibility that the behaviors reported in interviews may not be what actually happened.

Scope of data gathered

The implementation of the research design for this study was as planned, except for two items. Part of the original research design included analysis of a single repertory grid filled out by all participants. This analysis was not completed because only 16 of the 23 participants completed this task. In order to compensate for this issue, the research design was modified to add additional Leximancer data analysis. Secondly, a significant number of examples of shared artifacts were expected, but participants provided only fewer than five. The researcher reviewed the provided set, and correlated observations with interview feedback.

The data collection timeframe was for a finite portion of the overall project lifecycle. It is possible that an increased timeframe would have produced richer research materials, and thus enabled additional insights and findings. This, however, was outside the scope of the study parameters.

The researcher did not collect demographic data such as gender, age, and tenure with the company, for concern of preserving the anonymity of the participants. In a workgroup of the size interviewed (23 participants), the disclosure of these elements might make it possible to infer their identity and consequently compromise their anonymity. This was not a significant issue for the study since these individual characteristics were not in focus in the analysis, however, readers might expect this information to be made available.

The interview questions and process

The interview questions were minimalist to start, just the simple question: "what is shared?" This was received as an unexpected approach by the participants, and somewhat abrupt, and some people did have trouble getting started. In those cases, the interviewer provided limited additional description. The repertory grid elicitation segment of the characteristics of what was shared (grouping 2 items and a third contrasting one) was also awkward for some and required a facilitated example to get things going. The repertory grid rating process caused some additional questions for some participants, and six of the 23 did not complete this task.

Communication issues stemming from the interviews being conducted over the telephone, or even through face-to-face conversations, are another area of limitation. In order to address this issue, the participants were provided a copy of the interview transcript and had the opportunity to correct any errors or add any clarifications. A number of follow-ups occurred over the study period.

Anomalies in data and data analysis

As mentioned earlier, analysis could not be completed on a repertory grid, as an insufficient number was completed by participants, and on shared artifacts, as only a handful of examples were provided. Additional Leximancer 4 automated concept and thematic analysis was added to compensate for these omissions.

Dependability of research results through mechanisms of triangulation.

Triangulation of sources occurred through interviews with multiple members of the iProject workgroup, and method triangulation through both activity system analysis and modeling and Leximancer 4 concept and thematic analysis. The multi-voice data collection provided multiple vantage points on the overall context, with some participants reporting the same observations, while other varied observations demonstrated the individuality of context.

Finally, it is important to summarize and comment on what may be obvious: this is a human research activity; in a finite timeframe; with a limited set of focal points; in a single setting (albeit a complex one); with a set of people affiliated with one project; and in an industrial setting. This is both a great strength and a potential weakness. The opportunity to gather and analyse rich data, examining multiple vantage points of a situation, is a wonderful opportunity. The researcher made a conscious effort not to over-generalize the findings outside the scope within which they may apply; however, with the prevalence of distance collaboration and the ubiquity of information technologies for sharing information, the applicability of the findings have high potential.

This study provides a richly textured and often nuanced picture of the varying perspectives on the same set of information sharing mechanisms, and how workgroup members saw their characteristics relative to sharing in a data analytics software project. Expression of unique opinions, and the breadth of the human intellectual viewpoint emerges from the interview data, along with the rationale supporting those opinions. Moreover, although no two groups of this sort are identical, it is likely that the disciplinary and experiential diversity and viewpoints of this group of people are similar to other working groups, making the heterogeneous sensibility and awareness relevant to other settings.

5.6.1 Credibility, transferability, dependability, and confirmability

Lincoln and Guba (1985) stress the importance of credibility, confirmability, transferability, and dependability in assessing the trustworthiness of a research study. The research design for this study provided strengths for the trustworthiness and rigor of this qualitative study. Following are some commentaries about these characteristics in this study.

Credibility. This study engaged with the study participants over an 8-month period, with interviews phased over that period, augmented by preliminary planning discussions with project leaders at an earlier time, and opportunities for continued contact with the participants after the interviews concluded. This enabled a long period of interaction with the participants. "Prolonged engagement" is a factor of rigor and trustworthiness (Lincoln & Guba, 1985; Shenton, 2004; among others). A second factor adding to the credibility of this study is the multiple source triangulation in the collection of data (Lincoln & Guba, 1985). The 23 participants each provided their own perspective on the sharing prompt, and their view of the characteristics of those mechanisms, events, and artifacts. This multi-source data collection provided confirmation of the phenomena reported, and some healthy contrasts as well. The differing viewpoints suggest the candid nature of the interviews.

Transferability. Chapters 4 and 5 (Findings and Discussion) contain extensive interview quotes and details about the data collected. Thick descriptions are cited (Lincoln & Guba, 1985; Creswell & Miller, 2000) as a way for readers to understand as much as possible about the context of the interview comments and the environment of this study in order to form a comparison with environments they know.

Confirmability and Dependability. Due to the company business environment, business confidentiality and participant privacy protection requirements, it was not possible to perform any auditing or provide raw data from this study. The researcher acknowledges that, as mentioned in the previous section, this is a single case and does not make general claims about the repeatability of the findings in other environments. Additional studies are needed to validate these findings and conclusions.

5.7 SUMMARY

This chapter discussed the meaning of the study findings in the context of previous literature, highlighting confirmations of those studies as well as differences and new contributions to the literature. A new model of core information sharing distance and a metric for measuring the information sharing distance developed in this study are discussed in detail. The study contributed empirical data in two scant areas of information sharing research (information sharing in work activities with heterogeneous project roles, and with project membership changes). The chapter concludes with a review of a methodology innovation used in this study (activity theory, the two-phase repertory grid interview protocol, and Leximancer 4 data analysis), the limitations of this study, and aspects supporting credibility and transferability.

Chapter 6: Conclusion

6.1 CHAPTER OVERVIEW

The purpose of this chapter is to provide closure to the document, a summation of the key points of the dissertation work, a succinct statement of the answers to the research questions, and a perspective about potential future work to further our understanding of information sharing and collaboration over distance.

6.2 DISSERTATION SUMMARY

1. The overall research aim for this study was to gain insights on what people do when they share information in a software engineering workgroup, how they think about sharing, what they share, what others share with them, and what this all means for collaboration.
2. The literature foundation for information sharing spanned multiple disciplines: information behavior in the library and information science discipline, computer supported cooperative work in the computer science discipline, management, project management, and team development in the business discipline. This provides a broad landscape to view specific and contrasting perspectives.
3. The research design emphasized an approach that respected and preserved the individual and collective contexts of the study participants, starting with the repertory grid elicitation through the activity theory analysis and Leximancer 4 text analytics.
4. The analysis showed an interacting set of collaborators with their mechanisms of information sharing over a set of moments in time, and their viewpoints about those experiences and the mechanisms.
5. Multiple findings from this study are summarized below:
    a. This study showed that distance creates gaps that affect both information sharing and collaboration, in the dimensions of geography, time zone, discipline, role, project tenure, role, and organization. These were due to multiple factors, ranging from tacit knowledge, to differences in vocabulary, to technology

challenges affecting the fidelity and experience of mediated communication.

b. A wide variety of information sharing mechanisms are utilized by a workgroup with multi-dimensional distance factors, and a many of them are technology-mediated to a lesser or a greater extent. Pragmatism often wins out: participants choose mechanisms that work and have a good user experience.

c. A striking finding is the separation of shared information artifacts into two categories: (1) actual work exemplars, and (2) artifacts produced for the express purpose of sharing.

Actual work exemplars are produced in the course of performing the work itself, such as software code or a project plan. Because of their close proximity to the actual work, these artifacts may result in greater variation in understanding across the distances of the project team, particularly in the dimensions of discipline and role. These artifacts may carry higher "fidelity" to the workgroup members closest to those work activities, but at the same time may convey much less meaning across other dimensions of distance.

Specially produced derivatives are a contrasting category of artifacts. These are created for the express purpose of sharing information within the workgroup, or outside of it. Examples include artifacts such as a project status presentation, which highlights the progress toward milestones or deadlines, or a project overview PowerPoint, which reviews the purpose, objectives, and technologies of a project.

Each category of artifacts has advantages and disadvantages in the information sharing processes of a workgroup. The actual work exemplars have the advantage of requiring no additional effort to create them, and they contribute veracity about the project work itself. The derivative artifacts, on the other hand, require additional resources to create them, but can be tailored for the specific information sharing purpose and audience. Future investigations examining this division may yield helpful insights.

d. Workgroup participants often do not acknowledge or compensate for their distance factors, but address them in an emergent, evolutionary way through the actual work, and in fact may not understand that they exist. Examples of this include vocabulary, styles, and structures of meetings and communication.

e. Created in this study are two new models to frame a particular way of thinking about the phenomena under consideration, and the associated findings. They show two different representations of the multi-dimensional distance. The first, Information Sharing Discrete Distance Metric (Figure 4-3), illustrates these six factors as cumulative distance over which information sharing and collaboration must traverse. The second, Information Sharing Distance Model (Figure 5-1), portrays the six dimensions of distance (geography, time zone, discipline, role, project tenure, role, and organization) as core factors, surrounded by other important factors which may be seen as more perceptual, or preferential, and which may vary depending on a specific interaction or circumstance.

6.3 FINAL DESCRIPTION OF THE ANSWERS TO THE RESEARCH QUESTIONS

The core question of this research is: "how does information sharing occur in the distance collaboration of virtual teams?" This study found that information sharing is a critical activity in the life of a virtual team (and in other collaborating workgroup settings as well) but it is multifaceted in its manifestation, and is highly contextual. Especially in the knowledge work of software engineering, where architecture and design concepts are often expressed first as concepts, the ability to share, interact with, and understand information is necessary and important. Surprisingly, although the initial focus of the research was on virtual teams and geographical distance, other forms of distance emerged from the study as important, too.

Five subordinate questions explore more detailed aspects of this phenomena.

1. How do information sharing activities manifest themselves in distance collaboration?

   Information sharing activities in distance collaboration -- across the multi-dimensional distances of geography, time zone, discipline, role, project tenure, role, and organization – permeated the performing of the software engineering work of the workgroups studied. The sharing of information crossed essentially every line: formal/informal; spontaneous/planned; proactive/in response to a request; synchronous/asynchronous, direct products of the work/derivatives for sharing, and location. The performing of work requires information sharing.

2. When and in what kinds of circumstances does information sharing occur in distance collaboration?

   The circumstances of information sharing did have some differences in the lifecycle of a project. In the early work of a project, a turn-taking process of introduction of people, expertise and technologies served to build a knowledge foundation. As the project progressed, more targeted and purposeful sharing tied to specific tasks or milestones emerged. Temporal aspects also had some differences, as both co-located and remote workgroup members utilized synchronous mechanism when under deadline (face to face, or instant messaging), but in contrast used asynchronous mechanisms at other times.

3. What types of information sharing behaviors and forms of shared information can be identified?

   There were some commonly practiced information behaviors and patterns, as well as standard categories of forms of information sharing. Common interactive information sharing behaviors included leading/participating in meetings, presentations, demos, discussions, chat, and screen sharing. Categories of asynchronous communications and written expression included email, PowerPoint slides, chart decks, and contributing to/using information repositories (e.g., wikis). There were two categories of forms that information sharing divided into: events and artifacts. Some expressed confusion about the meaning of demos, prototypes, and mock-ups. The population of people receiving these expressions of abstract ideas often did not understand them.

4. What attributes are related to different types and forms of information sharing in distance collaboration?

The objective or goal of the individual influenced the different types and forms of information sharing. Participants carefully weighed the affordances, advantages, and disadvantages when choosing a particular information sharing approach, such as utilizing a set of architectural sketches to demonstrate design ideas. This choice favors a low-effort, lightweight nature versus coding a mock-up, which is much more labor intensive. Sometimes information sharing needed to be done on a point-to-point basis – person to person – instead of keeping all of the workgroup participants in the loop. However, these viewpoints varied widely, as there were many variations in perception across the workgroup. This is a deeply contextual, sometimes individual, perspective and judgment, and not broadly applicable even within the same workgroup.

5. What purposes does information sharing serve in distance collaboration?

The purposes served by information sharing in distance collaboration were also closely related to the objective or goal of the individual, and deeply contextual. Sometimes information sharing advanced the work a step or small way toward a bigger objective. Information sharing often served to clarify or confirm, to explore ideas, and to document understanding.

In summary, this study formed a detailed picture of a rich and varied human activity with many nuanced and interesting information sharing behaviors.

PRAGMATIC RECOMMENDATIONS

6.3.1 Insights from activity system contradictions

Activity system contradictions, as noted in section 4.5.2, provide energy for change, growth, development, and innovation. Changes, tensions, and conflicts affect multiple components in the activity system differently, and different activity systems differently (e.g., highest project level vs. individual person activity system). How these changes, tensions, and conflicts are managed determines if they are generative or constraining, at the project-level activity system. Lower level activity systems may not experience them in the short term as generative, for example, if the change requires creating new code and discarding the old code. Activity theory provides a structure for looking at the interrelationships between activity system components, between activity systems, and the opportunity to manage important contradictions in a direction that is generative and to evolve the activity in a positive direction.

There is an interdependency – an intertwining – between information sharing and collaboration which requires an adjustment for distance. There are multiple approaches to accomplish this, and the nature of the context – including factors such as the activity, organization(s), and overall environment -- is important to consider. However, the approach should include an assessment and acknowledgement of the distances that may exist. There needs to be an identification and understanding of where distances exist, where there is common ground, and where there are gaps. This awareness may help to close or reduce any gaps. The identification of a core set of information sharing discrete distance factors, using the models developed in this study, is useful for workgroups. This enables acknowledgment and accommodation of the distance(s). Although there are other important factors besides distance (such as culture and trust), a discussion of the six core distance dimensions in the workgroup helps build up mutual understanding and identify areas of uncommon ground.

Workgroups can also make adjustments in their information sharing mechanisms. There may be an opportunity for a workgroup to diagnose the dimensions of distance in a particular work activity, and to have explicit discussions about the purpose and practices of information sharing, and to develop some approaches that fit the composition of the workgroup and its members' needs and desires. This has the possibility to reduce the burden and noise of information sharing, and to improve the effectiveness of the information environment, the work activities, and associated object and outcome. It is also possible to make explicit the

Distance Type	Accommodation Strategies
Geographical	Technology: telephone conversations, conference calls, web conferences, screen-sharing, instant messaging, email, extension points in computer code (loosely coupled). Process: define work in greater detail, minimize dependencies, foster connection points.
Time Zone	Technology: Utilize extended hour telephone conversations, extended hour instant messaging, email. Process: define work in greater detail, adjust working hours.
Discipline	Process: Informal questions and answers, specialization.
Role	Process: Informal questions and answers.
Organization	Process: Identification of connection points, Informal questions and answers.
Project Tenure	Process and technology: Personal overview by longer-tenure workgroup member; questions/answers with longer-tenure workgroup member, use of documentation (e.g., wikis) to provide detailed project history.
Overall	Process: Discuss distance(s) explicitly; define vocabulary; customize information sharing practices for the workgroup.

Table 6-1: Practical recommendations to address information sharing distance

knowledge or information that had been tacit or implicit through shared information. This could involve approaches such as making modifications in information sharing mechanisms prior to just repurposing without changes. Alternatively, once understanding the existing distances, the workgroup could improve the situation by adjusting the collaboration itself to a loosely coupled approach, dividing the work with these gaps in mind, and/or assigning tasks with clean boundaries in mind.

6.4 FOR THE FUTURE

Looking ahead to the future, this research opens up at least three high potential pathways for follow-on investigations to shed more light on the details of information sharing behavior. These are: (1) issues of interest to investigate but not included in the scope this research due to limited time and/or data, (2) new ideas emerging from the interviews of participants, e.g., “unexpected” findings, and (3) directions identified by reflecting on the research findings, e.g., inconclusive findings and findings in conflict.

The primary issue of interest not included in this research is additional research with a project workgroup with the opportunity to work with completed repertory grids, either qualitatively or quantitatively, or both. This would provide the ability to compare the perspectives of each participant against the other in greater depth. Second is having access to a large number of shared information (artifacts such as presentations, architectural documents, and meeting transcripts), and to compare and contrast insights from one or both of these additional data sources with the findings from this study. Another research topic is an examination focusing on the creation stage of information to be shared and the variations and potential patterns that exist: looking at information that is individually created and used together; or created together and used together, or created together and used individually. Finally, the validation of this research through additional activity theory studies in different settings and contexts would be very helpful.

The second area -- emergent ideas from the participants and unexpected findings from the study -- includes further exploration of activity system contradictions that highlight distances over which information needs to travel, and further characterization of the distance dimensions such as project role and project tenure. A quantitative study looking at volume of different information sharing mechanisms would be interesting, and additional analysis about the variations across the same mechanism (e.g., different forms of "meeting") as well. Another research idea is the analysis of information sharing mechanisms (artifacts and events) in light of boundary crossing and within the boundary object framework, particularly in light of the prevalence of cross-disciplinary teams in the technical environments. An examination of sharing behaviors around *created* information (not *found* information) would be interesting. Much of the existing literature focuses on the seeking-sharing paradigm, but less on the creating-sharing activities. Another idea is additional exploration of the conceptual distance of specific information sharing mechanisms to the work itself. The study provided findings for this potential seventh information sharing distance dimension, and an opening for further exploration of this dimension. Also interesting would be a comparative study looking at volume of different information sharing mechanisms; and the study of a different company/group as a comparison to this case. Further explorations of the differences between shared information artifacts from the work itself versus artifacts created specifically for the purpose of sharing might be a fruitful direction investigate.

Finally, in the area of inconclusive or conflicting findings is the benefit of a tightly or loosely-coupled workgroup configuration. A detailed investigation is needed to understand more about the particular circumstances where one structure is preferable over another, and it would be beneficial to workgroups across many contexts.

It is clear from this study that information sharing is a vital and integral aspect of the social dimension of workgroups, particularly collaboration. It is also important to note that continuing research in this area to build a greater understanding of information sharing is important, from both theoretical and practical standpoints.

References

Ackerman, M. S., Dachtera, J., Pipek, V., & Wulf, V. (2013). Sharing knowledge and expertise: The CSCW view of knowledge management. *Computer Supported Cooperative Work (CSCW)*, *22*(4-6), 531–573.

Acquisti, A., & Gross, R. (2006, January). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Privacy enhancing technologies* (pp. 36-58). Springer Berlin Heidelberg.

Adams-Webber, J. (2006). A Review of: "Repertory Grid Theory, Research, and Applications." *Journal of Constructivist Psychology*, *19*(4), 351–353. http://doi.org/10.1080/13854040600689133

Aedo, I., Díaz, P., Carroll, J. M., Convertino, G., & Rosson, M. B. (2010). End-user oriented strategies to facilitate multi-organizational adoption of emergency management information systems. *Information processing & management*,*46*(1), 11-21.

Agarwal, N., & Rathod, U. (2006). Defining 'success' for software projects: An exploratory revelation. *International Journal of Project Management*, *24*(4), 358–370.

Agrawal, R., Evfimievski, A., & Srikant, R. (2003, June). Information sharing across private databases. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data* (pp. 86–97). ACM.

Ahituv, N., & Neumann, S. (1986). *Principles of Information Systems for Management (2nd Ed.)*. Dubuque, IA, USA: William C. Brown Publishers.

Akoumianakis, D. (2014a). Ambient affiliates in virtual cross-organizational tourism alliances: A case study of collaborative new product development. *Computers in Human Behavior*, *30*, 773–786.

Akoumianakis, D. (2014b). Boundary spanning tactics and "traceable" connections in cross-organizational virtual alliances: A case study. *Journal of Enterprise Information Management*, *27*(2), 197–227.

Algon, J. (1999). *The effect of task on the information-related behaviors of individuals in a work-group environment* (Order No. 9918306). Available from ProQuest Dissertations & Theses Global. (304533968). Retrieved from http://search.proquest.com/docview/304533968?accountid=13380

Allard, S., Levine, K. J., & Tenopir, C. (2009). Design engineers and technical professionals at work: Observing information usage in the workplace. *Journal of the American Society for Information Science and Technology*, *60*(3), 443-454.

Allen, D. (2011). Information behavior and decision making in time-constrained practice: A dual-processing perspective. *Journal of the American Society for Information Science and Technology*, *62*(11), 2165–2181. http://doi.org/10.1002/asi.21601

Allen, D., Karanasios, S., & Slavova, M. (2011). Working with activity theory: Context, technology, and information behavior. *Journal of the American Society for Information Science and Technology*, *62*(4), 776–788. doi:10.1002/asi.21441

Allen, T. J. (1977). Managing the flow of technology: Technology transfer and the dissemination of technological information within the research and development organization. *Boston, Mass.: Massachusetts Institute of Technology*.

Allen, T. J., & Henn, G.W. (2007). The Organization and Architecture of Innovation: Managing the Flow of Technology. Elsevier, Burlington, MA.

Ambler, S. W. (2014, February 4). The Non-Existent Software Crisis: Debunking the Chaos Report. Retrieved March 19, 2015, from http://www.drdobbs.com/architecture-and-design/the-non-existent-software-crisis-debunki/240165910

Artifact, [Def. 1]. (n.d.) in *OED Online*. Oxford University Press. Retrieved September 5, 2015 from http://www.oed.com.ezp01.library.qut.edu.au/view/Entry/11133?redirectedFrom=artifact#eidWeb.

Avis, J. (2009). Transformation or transformism: Engeström's version of activity theory? *Educational Review*, *61*(2), 151-165. doi:10.1080/00131910902844754

Baker, S. E., & Edwards, R. (2012, March 26). *How many qualitative interviews is enough* (Working Paper). NCRM. Retrieved from http://eprints.ncrm.ac.uk/2273/

Barab, S. A., Evans, M. A., & Baek, E. O. (2004). Activity theory as a lens for characterizing the participatory unit. In D. H. Jonassen (Ed.) *Handbook of research on educational communications and technology* (pp 199–214). Mahwah: Lawrence Erlbaum Associates.

Barrett, S., & Konsynski, B. (1982). Inter-organization information sharing systems. *MIS Quarterly*, *6(4)*, 93-105.

Barthelmess, P., & Anderson, K. M. (2002). A view of software development environments based on activity theory. *Computer Supported Cooperative Work (CSCW)*, *11*(1-2), 13-37.

Barua, A., & Ravindran, S. (1996). Reengineering information sharing behaviour in organizations. *Journal of Information Technology, 11*(3), 261–272.

Bassil, Y. (2012). A simulation model for the waterfall software development life cycle. *International Journal of Engineering & Technology,* 2(5), 742-749.

Bates, M. J. (2010). Information behaviour. In Bates, M.J. and Maack, M.N. (Eds), *Encyclopedia of Library and Information Sciences, 3rd ed.* (pp. 2381-2391). New York: CRC Press.

Bateson, G. (1972). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. New York: Chandler publications.

Bayerl, P. S., & Lauche, K. (2008). Coordinating high-interdependency tasks in asymmetric distributed teams. In *Proceedings of the ACM 2008 conference on Computer-supported cooperative work* (pp. 417-426). ACM. doi:10.1145/1460563.1460631

Beecham, S., Noll, J., & Richardson, I. (2014, August). Using agile practices to solve global software development problems–a case study. In *Global Software Engineeering Workshops (ICGSEW), 2014 IEEE International Conference on* (pp. 5–10). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6912078

Belkhatir, N., & Ahmed-Nacer, M. (1995). Major issues on process software engineering environments. *Information Sciences—Informatics and Computer Science: An International Journal*, *83*(1-2), 1-21.

Benford, S., Greenhalgh, C., Rodden, T., & Pycock, J. (2001). Collaborative virtual environments. *Communications of the Association for Computing Machinery, 44*(7), 79- 85. doi:10.1145/379300.379322

Bentley, R., Horstmann, T., Sikkel, K., & Trevor, J. (1995). Supporting collaborative information sharing with the World Wide Web: The BSCW shared workspace system. In *Proceedings of the 4th International WWW Conference* (1), 63–74.

Bjørn, P., Esbensen, M., Jensen, R. E., & Matthiesen, S. (2014). Does distance still matter? Revisiting the CSCW fundamentals on distributed collaboration. *ACM Transactions on Computer-Human Interaction, 21*(5), 1-27. http://doi.org/10.1145/2670534

Blomberg, J. (2008). Negotiating meaning of shared information in service system encounters. *European Management Journal, 26*(4), 213–222.

Bock, G. W., & Kim, Y. G. (2002). Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing. *Information Resources Management Journal*, *15*(2), 14–21.

Bødker, S. (1991). Activity theory as a challenge to system design. In Nissen, H.E., Klein, H.K., and Hirschheim, R. (Eds*), Information System Research: Contemporary Approaches and Emergent Traditions* (pp. 551–564). North Holland: Elsevier Science Publisher BV.

Boehm, B. (2002). Get ready for agile methods, with care. *Computer*, *35*(1), 64–69.

Boer, N. I., Van Baalen, P. J., & Kumar, K. (2002, January). An activity theory approach for studying the situatedness of knowledge sharing. In *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on* (pp. 1483-1492). IEEE.

Bonebright, D. A. (2010). 40 years of storming: a historical review of Tuckman's model of small group development. *Human Resource Development International*, *13*(1), 111–120.

Braun, S., Hefke, M., Schmidt, A., & Sevilmis, N. (2007, May). Im Wissensnetz: Linked information processes in research networks. In *Proceedings of the German e-Science Conference 2007,* Max Planck Digital Library, ID 316584.0.

Bruce, H., Fidel, R., Pejtersen, A., Dumais, S., Grudin, J., & Poltrock, S. (2003). A comparison of the collaborative information retrieval behaviour of two design teams. *The New Review of Information Behaviour Research*, *4*(1), 139-153. doi:10.1080/14716310310001631499

Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science*, *42*(5), 351–360.

Bunderson, J. S. (2003). Team member functional background and involvement in management teams: Direct effects and the moderating role of power centralization. *Academy of Management Journal*, *46*(4), 458–474.

Bunderson, J. S., & Sutcliffe, K. M. (2002). Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Academy of Management Journal*, *45*(5), 875–893.

Capurro, R., & Hjørland, B. (2003). The concept of information. *Annual Review of Information Science and Technology, 37*(1), 343-411. doi:10.1002/aris.1440370109

Carmel, E. & Abbott, P. (2007). Why 'nearshore' means that distance matters. *Communications of the ACM*, *50*(10), 40-46.

Carmel, E., & Agarwal, R. (2001). Tactical approaches for alleviating distance in global software development. *Institute of Electrical and Electronics Engineers (IEEE) Software*, *18*(2), 22–29.

Case, D. O. (2012). *Looking for information: A survey of research on information seeking, needs and behavior*. Emerald Group Publishing.

Chatman, E. A. (1996). The impoverished life-world of outsiders. *Journal of the American Society for Information Science and Technology*, *47*(3), 193–206.

Cheng, L., Souza, C. R. D., Hupfer, S., Patterson, J., & Ross, S. (2004). Building collaboration into IDEs. *Queue*, *1*(9), 40-50. doi:10.1145/966789.966803

Cho, H. (2008). Collaborative information seeking in intercultural computer-mediated communication groups: testing the influence of social context using social network analysis. *Communications Research*, *35*(4), 548-573.

Churchill, E. F., Nelson, L., & Denoue, L. (2003). Multimedia fliers: Information sharing with digital community bulletin boards. In *Communities and technologies* : *proceedings of the first International Conference on Communities and Technologies, C&T 2003* (pp. 97-117). Springer Netherlands.

Clark, H. H. (1996) *Using language.* Cambridge: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, *13*(1991), 127–149.

Cole, M. (1993, March/April). Vygotsky and writing: Reflections from a distant discipline. Paper presented at the annual convention of the conference on College Composition and Communication, San Diego. Retrieved September 14, 2015, from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED360627

Cole, M. (1999). Cultural psychology: Some general principles and a concrete example. In Engeström, Y., Miettinen, R., & Punamäki-Gitai, R.-L. (eds.), *Perspectives on activity theory* (pp.87-106). Cambridge University Press.

Cole, M., Göncü, A., & Vadeboncoeur, J. A. (2014). Experience, imagination, and action: Versions of artifact mediation. *Mind, Culture, and Activity*, *21*(4), 275–277. http://doi.org/10.1080/10749039.2014.958929

Collaboration [Def. 1] .(n.d.) in OED Online. Oxford University Press. Retrieved March 28, 2011 from http://www.oed.com.ezp01.library.qut.edu.au/viewdictionaryentry/Entry/36197

Collins, P., Shukla, S., & Redmiles, D. (2002). Activity theory and system design: A view from the trenches. *Computer Supported Cooperative Work* (CSCW), *11*(1-2), 55-80.

Connelly, C. E., & Kelloway, E. K. (2003). Predictors of employees' perceptions of knowledge sharing cultures. *Leadership & Organization Development Journal*, *24*(5), 294–301.

Constant, D., Kiesler, S., & Sproull, L. (1994). What's mine is ours, or is it? A study of attitudes about information sharing. *Information systems research*, *5*(4), 400–421.

Constant, D., Sproull, L., & Kiesler, S. (1996). The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization Science*, *7*(2), 119–135.

Cool, C., & Xie, H. (2000). Patterns of information use, avoidance and evaluation in a corporate engineering environment. In *Proceedings of the ASIS Annual Meeting* (Vol. 37, pp. 462–472).

Cramton, C. (2001). The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration. *Organization Science*, *12*(3), 346-371.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice*, *39*(3), 124-130.

Croker, A., Higgs, J., & Trede, F. (2009). What do we mean by 'collaboration' and when is a 'team' not a 'team'?: A qualitative unbundling of terms and meanings. *Qualitative Research Journal*, *9*(1), 28-42.

Cronin, B. (1982). Invisible colleges and information transfer - A review and commentary with particular reference to the social sciences. *Journal of Documentation*, *38*(3), 212–236.

Cummings, J. N., & Kiesler, S. (2008, November). Who collaborates successfully?: Prior experience reduces collaboration barriers in distributed interdisciplinary research. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 437-446). ACM.doi:10.1145/1460563.1460633

Dalal, S., & Chhillar, R. S. (2012). Case Studies of Most Common and Severe Types of Software System Failure. *International Journal of Advanced Research in Computer Science and Software Engineering*, *2*(8), 341-347. Retrieved from http://csis.pace.edu/~marchese/SE616_New/L1/V2I800209.pdf

Davenport, E. (2002). Mundane knowledge management and microlevel organizational learning: An ethological approach. *Journal of the American Society for Information Science and Technology*, *53*(12), 1038–1046.

Davenport, E., & Hall, H. (2005). Organizational knowledge and communities of practice. *Annual Review of Information Science and Technology*, *36*(1), 170–227. http://doi.org/10.1002/aris.1440360105

Davydov, V. V. (1999). The content and unsolved problems of activity theory. In Engeström, Y., Miettinen, R., & Punamäki-Gitai, R.-L. (eds.), *Perspectives on activity theory* (pp. 39-52). Cambridge University Press.

Dearman, D., Kellar, M., & Truong, K. N. (2008, November). An examination of daily information needs and sharing opportunities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 679–688). ACM. http://doi.org/10.1145/1460563.1460668

Dervin, B. (1983, May). An overview of sense-making research: Concepts, methods and results. Paper presented at the annual meeting of the International Communication Association, Dallas, TX. Retrieved from: http://faculty.washington.edu/wpratt/MEBI598/Methods/An%20Overview%20of%20Sense-Making%20Research%201983a.htm

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. In Glazier, J.D., Powell, R.R., (eds.), *Qualitative research in information management* (pp. 61–84). Englewood, CO: Libraries Unlimited.

Dervin, B. (1999). On studying information seeking methodologically: the implications of connecting metatheory to method. *Information Processing and Management*, *35*(6), 727– 750.

Dixon, J. (2010, October). Pentaho, Hadoop, and Data Lakes. Retrieved from http://www.pentaho.com/blog/2010/10/15/pentaho-hadoop-and-data-lakes

Dybå, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. *Information and software technology*, *50*(9), 833–859.

Ehrlich, K., & Chang, K. (2006, October). Leveraging expertise in global software teams: Going outside boundaries. In *Global Software Engineering, 2006. ICGSE'06. International Conference on* (pp. 149-158). IEEE.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, *45*(3), 171–212.

Ellis, D., & Haugan, M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of documentation*, *53*(4), 384 - 403. doi:10.1108/EUM0000000007204

Engelbart, D. C., & English, W. K. (1968). A research center for augmenting human intellect. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I* (pp. 395–410). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1476645

Engeström, Y. (1987). *Learning by Expanding: An activity-theoretical approach to developmental research*. Helsinki, Finland: Orienta-Konsultit Oy.

Engeström, Y. (1999). Activity theory and individual and social transformation. In Engeström, Y., Miettinen, R., & Punamäki, R.-L. (Eds.). *Perspectives on activity theory* (pp. 19-38*).* Cambridge University Press.

Engeström, Y. (2000). From individual action to collective activity and back: developmental work research as an interventionist methodology. In Luff, P., Hindmarsh, J., & Heath, C., (Eds.), *Workplace Studies: Recovering Work Practice and Information System Design* (pp. 150-168). Cambridge: Cambridge University Press.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of education and work*, *14*(1), 133–156.

Engeström, Y. (2006). From well-bounded ethnographies to intervening in mycorrhizae activities. *Organization Studies*, *27*(12), 1783-1793.

Engeström, Y. (2007). From communities of practice to mycorrhizae. In Hughes, J., Jewson, N., and Unwin, L. (eds.*), Communities of practice: Critical perspectives* (pp. 41–54). London: Routledge. Retrieved from http://www.open.ac.uk/opencetl/files/opencetl/file/ecms/web-content/Engestrom-(2007)-From-communities-of-practice-to-mycorrhizae.pdf

Engeström, Y. (2008a, September). The future of activity theory: A rough draft. Presented at the International Society of Cultural-historical Activity Research Conference, San Diego, CA, USA. Retrieved from http://lchc.ucsd.edu/mca/Paper/ISCARkeyEngestrom.pdf

Engeström, Y. (2008b). *From teams to knots: Activity-theoretical studies of collaboration and learning at work*. Cambridge University Press.

Engeström, Y. (2009). The future of activity theory: A rough draft. In Sannino, A., Daniels, H., and Gutiérrez, K. (Eds.), *Learning and expanding with activity theory*. (pp.303-328). Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511809989.020

Engeström, Y., Kaatrakoski, H., Kaiponen, P., Lahikainen, J., Laitinen, A., Myllys, H., Rantavuori, J., Sinikara, K. (2012). Knotworking in Academic Libraries: Two Case Studies from the University of Helsinki. *LIBER Quarterly, 21*(3/4), 387–405.

Engeström, Y., Miettinen, R., & Punamäki-Gitai, R. L. (1999). *Perspectives on activity theory*. Cambridge University Press.

Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, *5*(1), 1–24. http://doi.org/10.1016/j.edurev.2009.12.002

Erdelez, S. (1997). Information encountering: a conceptual framework for accidental Erdelez, S. (1997, August). Information encountering: a conceptual framework for accidental information discovery. In *Proceedings of an international conference on Information seeking in context* (pp. 412-421). Taylor Graham Publishing.

Evans, B. M., & Chi, E. H. (2008, November). Towards a model of understanding social search. In *Proceedings of the Association for Computing Machinery 2008 conference on Computer-supported cooperative work*, 485–494. ACM.

Fidel, R., Bruce, H., Pejtersen, A., Dumais, S., Grudin, J., & Poltrock, S. (2000). Collaborative information retrieval. *The New Review of Information Behaviour Research*, *1*(1), 235– 247.

Fidel, R., & Green, M. (2004). The many faces of accessibility: engineers' perception of information sources. *Information Processing & Management*, *40*(3), 563–581.

Fidel, R., Mark Pejtersen, A., Cleal, B., & Bruce, H. (2004). A multidimensional approach to the study of human-information interaction: A case study of collaborative information retrieval. *Journal of the American Society for Information Science and Technology*, *55*(11), 939-953.

Fiol, C. M., & O'Connor, E. J. (2005). Identification in face-to-face, hybrid, and pure virtual teams: Untangling the contradictions. *Organization Science*, *16*(1), 19–32.

Fischer, G. (2005). Distances and diversity: sources for social creativity. In *Proceedings of the 5th Conference on Creativity and cognition*, 128-136. London, United Kingdom: ACM. doi:10.1145/1056224.1056243

Fisher, K. E., & Julien, H. (2009). Information behavior. *Annual Review of Information Science and Technology*, *43*(1), 1–73. http://doi.org/10.1002/aris.2009.1440430114

Fjuk, A., & Ludvigsen, S. (2001, March) The complexity of distributed collaborative learning: Unit of analysis. *First European Conference on Computer-Supported Collaborative Learning*. University of Maastricht. Retrieved from https://www.isls.org/cscl/Euro2001/Papers/51.doc

Foley, C., & Smeaton, A. F. (2009). Synchronous collaborative information retrieval: Techniques and evaluation. In *Advances in Information Retrieval* (pp. 42-53). Springer Berlin Heidelberg.

Foner, L. N. (1995, March). Clustering and information sharing in an ecology of cooperating agents. In *Working Notes of the AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments, Stanford University, Stanford, CA.*

Foot, K. A. (2014). Cultural-historical activity theory: Exploring a theory to inform practice and research. *Journal of Human Behavior in the Social Environment*, *24*(3), 329–347.

Fourie, I. (2013). Twenty-first century librarians: time for Zones of Intervention and Zones of Proximal Development? *Library Hi Tech*, *31*(1), 171–181. http://doi.org/10.1108/07378831311304001

Franz, T. M., & Larson, J. R. (2002). The impact of experts on information sharing during group discussion. *Small Group Research*, *33*(4), 383–411.

Gal-Or, E. (1985). Information Sharing in Oligopoly. *Econometrica*, *53*(2), 329-343.

Gilson, L. L., Lim, H. S., Luciano, M. M., & Choi, J. N. (2013). Unpacking the cross-level effects of tenure diversity, explicit knowledge, and knowledge sharing on individual creativity. *Journal of Occupational and Organizational Psychology*, *86*(2), 203–222.

Gilson, L. L., Maynard, M. T., Young, N. C. J., Vartiainen, M., & Hakonen, M. (2015). Virtual Teams Research 10 Years, 10 Themes, and 10 Opportunities. *Journal of Management*, *41*(5), 1313-1337.

Glaser, B. & Strauss, A. (1967). The discovery of grounded theory. *Aldin, Chicago*.

Goggins, S. P., Mascaro, C., & Valetto, G. (2013). Group informatics: A methodological approach and ontology for sociotechnical group research. *Journal of the American Society for Information Science and Technology*, *64*(3), 516–539. http://doi.org/10.1002/asi.22802

Goh, C. & Hooper, V. (2009). Knowledge and information sharing in a closed information environment. Journal of Knowledge Management, 13(2), pp. 21-34.

Golovchinsky, G., Qvarfordt, P., & Pickens, J. (2009). Collaborative Information Seeking. *Institute of Electrical and Electronics Engineers (IEEE) Computer*, *42*(3), 47-51.

Goudreau, J. (2013). Back to the stone age? New Yahoo CEO Marissa Mayer bans working from home. *Forbes*, *February 25*.

Graetz, K. A., Boyle, E. S., Kimble, C. E., Thompson, P., & Garloch, J. L. (1998). Information sharing in face-to-face, teleconferencing, and electronic chat groups. *Small Group Research*, *29*(6), 714–743.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 1360–1380.

Grubb, A. M., & Begel, A. (2012, February). On the perceived interdependence and information sharing inhibitions of enterprise software engineers. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 1337–1346). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2145403

Gutwin, C., Penner, R., & Schneider, K. (2004, November). Group awareness in distributed software development. In *Proceedings of the 2004 Association for Computing Machinery Conference on Computer-supported cooperative work*, 72–81. ACM.

Haeussler, C. (2011). Information-sharing in academia and the industry: A comparative study. *Research Policy*, *40*(1), 105–122.

Haeussler, C., Jiang, L., Thursby, J., & Thursby, M. (2014). Specific and general information sharing among competing academic researchers. *Research Policy*, *43*(3), 465–475. http://doi.org/10.1016/j.respol.2013.08.017

Hansen, P., & Järvelin, K. (2005). Collaborative information retrieval in an information-intensive domain. *Information Processing & Management*, *41*(5), 1101-1119. doi:10.1016/j.ipm.2004.04.016

Hassan Ibrahim, N., & Allen, D. (2012). Information sharing and trust during major incidents: Findings from the oil industry. *Journal of the American Society for Information Science and Technology*, *63*(10), 1916–1928.

Hendriks, P. (1999). Why Share Knowledge? The Iinfluence of ICT on the motivation for knowledge sharing. *Knowledge and Process Management*, *6*(2), 91–100.

Herbsleb, J. D. (2007, May). Global software engineering: The future of socio-technical coordination. In *2007 Future of Software Engineering* (pp. 188-198). IEEE Computer Society.

Hertzum, M. (2002). The importance of trust in software engineers' assessment and choice of information sources. *Information and Organization*, *12*(1), 1–18.

Hertzum, M. (2008). Collaborative information seeking: The combined activity of information seeking and collaborative grounding. *Information Processing and Management*, *44*(2), 957–962.

Hertzum, M., & Pejtersen, A. (2000). The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing & Management*, *36*(5), 761- 778. doi:10.1016/S0306-4573(00)00011-X

Hinds, P. J., & Weisband, S. P. (2003). Knowledge sharing and shared understanding in virtual teams. In Gibson, C.B., & Cohen, S.G., *Virtual teams that work: Creating conditions for virtual team effectiveness* (pp. 21–36). San Francisco: Jossey-Bass.

Hirsh, S., & Dinkelacker, J. (2004). Seeking information in order to produce information: An empirical study at Hewlett Packard Labs. *Journal of the American Society for Information Science and Technology*, *55*(9), 807–817. http://doi.org/10.1002/asi.20024

Holmström, H., Fitzgerald, B., Agerfalk, P. J., & Conchúir, E. Ó. (2006). Agile practices reduce distance in global software development. *Information Systems Management*, *23*(3), 7–18.

Huang, H., Stvilia, B., Jörgensen, C., & Bass, H. W. (2012). Prioritization of data quality dimensions and skills requirements in genome annotation work. *Journal of the American Society for Information Science and Technology*, *63*(1), 195–207. http://doi.org/10.1002/asi.21652

Humphrey, W. S. (1989). The software engineering process: definition and scope. *ACM Special Interest Group on Software (SIGSOFT) Software Engineering Notes*, *14*(4), 82-83. doi:10.1145/75111.75122

Hupfer, S., Cheng, L. T., Ross, S., & Patterson, J. (2004, November). Introducing collaboration into an application development environment. In *Proceedings of the 2004 Association for Computing Machinery (ACM) conference on Computer-supported cooperative work,* 21-24. Chicago, Illinois, USA: ACM. doi:10.1145/1031607.1031611

Hussey, T. (2005). Failed IT projects cost billions: Train wreck in slow motion. *Visionpace*. Retrieved January 11, 2010, from http://blog.visionpace.com/2005/11/failed it proje.html

Hutchins, E. (2000). Distributed cognition. *International encyclopedia of the social and behavioral sciences. Elsevier Science*. Retrieved from http://hip.cntb.webfactional.com/hip-site/uploads/2014/07/Hutchins_DistributedCognition.pdf

Isah, E. E., & Byström, K. (2015). Physicians' learning at work through everyday access to information. *Journal of the Association for Information Science and Technology*. http://doi.org/10.1002/asi.23378

Jangir, S. K., Gupta, N., & Agrawal, S. (2012). Evolution and melioration of software management processes. *International Journal of Software Engineering & Applications*, *3*(1), 61-83.

Jarvenpaa, S. L., & Staples, D. S. (2000). The use of collaborative electronic media for information sharing: an exploratory study of determinants. *The Journal of Strategic Information Systems*, *9*(2), 129–154.

Jones, R. L., Edwards, C., & Viotto Filho, I. T. (2014). Activity theory, complexity and sports coaching: an epistemology for a discipline. *Sport, education and society*, *0*(0), 1–17. http://doi.org/10.1080/13573322.2014.895713

Jorgensen, M. (2014). What we do and don't know about software development effort estimation. *IEEE Software*, *31*(2), 37–40. http://doi.org/10.1109/MS.2014.49

Kamal, M., Weerakkody, V., & Irani, Z. (2011). Analyzing the role of stakeholders in the adoption of technology integration solutions in UK local government: An exploratory study. *Government Information Quarterly, 28(2),* 200-210.

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. MIT Press.

Kaptelinin, V., Nardi, B., (2012). *Activity theory: A primer for HCI Research*. San Rafael: Morgan & Claypool Publishers. http://dx.doi.org/10.2200/S00413ED1V01Y201203HCI013

Kaptelinin, V., Nardi, B. A., & Macaulay, C. (1999). Methods & tools: The activity checklist: a tool for representing the "space" of context. *interactions, 6(4)*, 27–39.

Karunakaran, A., Reddy, M. C. ., & Spence, P. R. (2013). Toward a model of collaborative information behavior in organizations. *Journal of the American Society for Information Science & Technology*, *64*(12), 2437–2451. http://doi.org/10.1002/asi.22943

Kealey, T., & Ricketts, M. (2014). Modelling science as a contribution good. *Research Policy*, *43*(6), 1014–1024.

Kelly, G. A. (1955). *The psychology of personal constructs. Volume 1: A theory of personality*. WW Norton and Company.

Kim, K. Y., Manley, D. G., & Yang, H. (2006). Ontology-based assembly design and information sharing for collaborative product development. *Computer-Aided Design*, *38*(12), 1233–1250.

King, D. W., with Casto, J., & Jones, H. (1994). Communication by engineers: A literature review of engineers' information needs, seeking processes, and use. Washington, DC: Council on Library Resources. Retrieved from http://eric.ed.gov/?id=ED385295

Kleiner, M. M., & Bouillon, M. L. (1991). Information sharing of sensitive business data with employees. *Industrial Relations: A Journal of Economy and Society*, *30*(3), 480–491.

Ko, A. J., DeLine, R., & Venolia, G. (2007, May). Information needs in collocated software development teams. In *Proceedings of the 29th international conference on Software Engineering* (pp. 344-353). IEEE Computer Society. Retrieved from http://portal.acm.org/citation.cfm?id=1248820.1248867

Koppman, S., & Gupta, A. (2014). Navigating the mutual knowledge problem: a comparative case study of distributed work. *Information Technology & People*, *27*(1), 83–105.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science, 42*(5), 361–371.

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. In *Context and consciousness: activity theory and human-computer interaction* (pp. 17-44). MIT Press.

Langemeyer, I., & Roth, W. M. (2006). Is cultural-historical activity theory threatened to fall short of its own principles and possibilities as a dialectical social science?. *Outlines. Critical Practice Studies, 8*(2), 20-42.

Lantolf, J., & Thorne, S. (2007). Sociocultural theory and second language learning. In. van Patten, B, & Williams, J. (eds.), *Theories in Second Language Acquisition* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum.

Larsson, A. (2003, November). Making sense of collaboration: the challenge of thinking together in global design teams. In *Proceedings of the 2003 international ACM SIGGROUP conference on supporting group work* (pp. 153-160). ACM.. doi:10.1145/958160.958184

Lauche, K. (2005). Collaboration among designers: Analysing an activity for system development. *Computer Supported Cooperative Work (CSCW), 14*(3), 253-282.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lehtinen, T. O., Mäntylä, M. V., Vanhanen, J., Itkonen, J., & Lassenius, C. (2014). Perceived causes of software project failures – An analysis of their relationships. *Information and Software Technology*, *56*(6), 623–643. http://doi.org/10.1016/j.infsof.2014.01.015

Leont'v, A. N. (1978). *Activity, consciousness and personality*. Englewood Cliffs, NJ: Prentice Hall. (Hall, M.J translator).

Leont'v, A.N. (1979). The problem of activity in psychology. In Wertsch, J.V. (Ed.). *The concept of activity in Soviet psychology* (pp. 37-71). Armonk, NY: Sharpe.

Leont'v, A. N. (1981). *Problems of the development of the mind*. Moscow: Progress.

Linberg, K. R. (1999). Software developer perceptions about software project failure: a case study. *Journal of Systems & Software*, *49*(2-3), 177–192.

Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic Inquiry (Vol 75)*. Newbury Park, CA: Sage Publications.

Maglaughlin, K. L., & Sonnenwald, D. H. (2005, July). Factors that impact interdisciplinary natural science research collaboration in academia. In *Proceedings of the ISSI* (pp. 24-25).

Maslow, A. H. (1943). A theory of human motivation. *Psychological review,50*(4), 370.

May, L. J. (1998). Major causes of software project failures. *CrossTalk: The Journal of Defense Software Engineering*, *11*(6), 9–12.

Mervyn, K., & Allen, D. K. (2012). Sociospatial context and information behavior: Social exclusion and the influence of mobile information technology. *Journal of the American Society for Information Science and Technology*, *63*(6), 1125–1141. http://doi.org/10.1002/asi.22626

Mesmer-Magnus, J. R., & DeChurch, L. A. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology*, *94*(2), 535-546.

Mesmer-Magnus, J. R., DeChurch, L. A., Jimenez-Rodriguez, M., Wildman, J., & Shuffler, M. (2011). A meta-analytic investigation of virtuality and information sharing in teams. *Organizational Behavior and Human Decision Processes*, 115(2), 214-225.

Milewski, A. (2007). Global and task effects in information-seeking among software engineers. *Empirical Software Engineering*, *12*(3), 311-326.

Milewski, A., Tremaine, M., Egan, R., Zhang, S., Köbler, F., & O'Sullivan, P. (2007, October). Information bridging in a global organization. In *Proceedings of the 2007 conference of the center for advanced studies on Collaborative research* (pp. 346-355). Richmond Hill, Ontario, Canada: ACM. doi:10.1145/1321211.1321258

Millen, D. R., & Dray, S. M. (2000, June). Information sharing in an online community of journalists. In *Aslib Proceedings* (Vol. 52, pp. 166–173). MCB UP Ltd. Retrieved from http://www.emeraldinsight.com/doi/abs/10.1108/EUM0000000007011

Mills, A. J., Durepos, G., & Wiebe, E. (2009). *Encyclopedia of case study research* (Vol. 2). Sage Publications.

Miranda, S. M., & Saunders, C. S. (2003). The social construction of meaning: An alternative perspective on information sharing. *Information systems research*, *14*(1), 87–106.

Mishra, J. L., Allen, D. K., & Pearman, A. D. (2011, May). Activity theory as a methodological and analytical framework for information practices in emergency management. In *Proceedings of the 8th International ISCRAM Conference–Lisbon* (Vol. 1). Retrieved from http://www.iscramlive.org/ISCRAM2011/proceedings/papers/140.pdf

Mishra, J., Allen, D., & Pearman, A. (2015). Information seeking, use, and decision making. *Journal of the Association for Information Science and Technology, 66*(4), 662–673. http://doi.org/10.1002/asi.23204

Modi, S., Abbott, P., & Counsell, S. (2013, August). Negotiating common ground in distributed agile development: A case study perspective. In *Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on* (pp. 80–89). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6613071

Moye, N. A., & Langfred, C. W. (2004). Information sharing and group conflict: Going beyond decision making to understand the effects of information sharing on group performance. *International Journal of Conflict Management*, *15*(4), 381–410.

Mwanza, D. (2001, July). Where theory meets practice: A case for an activity theory based methodology to guide computer system design. In *Proceedings of INTERACT' 2001: Eighth IFIP TC 13 Conference on Human-Computer Interaction* (Vol. 1, pp. 342-349).

Nardi, B. A. (1996). *Context and consciousness: activity theory and human-computer interaction*. Boston: MIT Press.

Nardi, B. A. (1998). Activity theory and its use within human-computer interaction. *The Journal of the Learning Sciences*, *7*(2), 257–261. http://doi.org/10.1207/s15327809jls0702_6

Nardi, B., & Harris, J. (2006, November). Strangers and friends: collaborative play in World of Warcraft. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 149-158). Banff, Alberta, Canada: ACM. doi:10.1145/1180875.1180898

Nissen, H. A., Evald, M. R., & Clarke, A. H. (2014). Knowledge sharing in heterogeneous teams through collaboration and cooperation: Exemplified through public–private-innovation partnerships. *Industrial Marketing Management*, *43*(3), 473–482.

Ocker, R. J., & Hiltz, S. R. (2012, January). Learning to work in partially distributed teams: The impact of team interaction on learning outcomes. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 88–97). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6148618

Oldham, G. R., & Da Silva, N. (2015). The impact of digital technology on the generation and implementation of creative ideas in the workplace. *Computers in Human Behavior*, *42*, 5–11. http://doi.org/10.1016/j.chb.2013.10.041

Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-computer interaction, 15*(2), 139–178.

Pallot, M., Martínez-Carreras, M. A., & Prinz, W. (2010). Collaborative distance: A framework for distance factors affecting the performance of distributed collaboration. *International Journal of e-Collaboration*, *6*(2), 1–32. http://doi.org/10.4018/jec.2010040101

Pan, S. L., & Leidner, D. E. (2003). Bridging communities of practice with information technology in pursuit of global knowledge sharing. *The Journal of Strategic Information Systems*, *12*(1), 71–88.

Pattinson, S., & Preece, D. (2014). Communities of practice, knowledge acquisition and innovation: A case study of science-based SMEs. *Journal of Knowledge Management*, *18*(1), 107–120.

Pentland, A. (2014). *Social physics: How good ideas spread - the lessons from a new science*. Penguin.

Pilerot, O. (2012). LIS research on information sharing activities – people, places, or information. *Journal of Documentation*, *68*(4), 559–581. http://doi.org/10.1108/00220411211239110

Pilerot, O. (2014). Making design researchers' information sharing visible through material objects. *Journal of the Association for Information Science and Technology*, *65*(10), 2006-2016. http://doi.org/10.1002/asi.23108

Pilerot, O., & Limberg, L. (2011). Information sharing as a means to reach collective understanding: A study of design scholars' information practices. *Journal of Documentation*, *67*(2), 312–333.

Poltrock, S., Grudin, J., Dumais, S., Fidel, R., Bruce, H., & Pejtersen, A. M. (2003, November). Information seeking and sharing in design teams. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work* (pp. 239-247). ACM.

Powell, A., Piccoli, G., & Ives, B. (2004). Virtual teams: a review of current literature and directions for future research. *ACM Database for Advances in Information Systems*, *35*(1), 6-36.

Prekop, P. (2002). A qualitative study of collaborative information seeking. *Journal of Documentation*, *58*(5), 533 - 547. doi:10.1108/00220410210441000

Price, D. D. S. (1963). Big science, little science. New York: *Columbia University Press*.

Rafaeli, S., & Raban, D. R. (2005). Information sharing online: a research challenge. *International Journal of Knowledge and Learning*, *1*(1), 62–79.

Rafaeli, S., & Ravid, G. (2003). Information sharing as enabler for the virtual team: An experimental approach to assessing the role of electronic mail in disintermediation. *Information Systems Journal*, *13*(2), 191–206.

Razavi, M. N., & Iverson, L. (2006, November). A grounded theory of information sharing behavior in a personal learning space. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 459–468). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1180946

Reddy, M. C., Dourish, P., & Pratt, W. (2001, January). Coordinating heterogeneous work: Information and representation in medical care. In *ECSCW 2001* (pp. 239–258). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/0-306-48019-0_13.

Reijonen, I. & Talja, S. (2006). Information sharing practices in multidisciplinary research groups. In Guerrero-Bote, V.P. (ed) *Current Research in Information Sciences and Technologies: Multidisciplinary Approaches to Global Information Systems, Vol. 1*. (pp. 461-465). Bajadoz: Open Institute of Knowledge.

Ren, Y., Sha, C., Qian, W., Zhou, A., Ooi, B. C., & Tan, K. L. (2003, May). Explore the small world phenomena in pure P2P information sharing systems. *In Cluster Computing and the Grid, 2003. Proceedings. CCGrid 2003. 3rd IEEE/ACM International Symposium on* (pp. 232-239). IEEE.

Rioux, K. S. (2004*). Information acquiring-and-sharing in Internet-based environments: An exploratory study of individual user behaviors.* (Unpublished dissertation) The University of Texas at Austin, Texas Retrieved from http://www.library.utexas.edu/etd/d/2004/riouxks042/riouxks042.pdf

Rioux, K. S. (2005). Information Acquiring-and- Sharing. In S. Erdelez, S., Fisher, K & McKechnie, L. (Eds.), *Theories of Information Behavior: A Researcher's Guide* (pp.169-173). Medford, NJ: Information Today, Inc.

Robinson, M. A. (2010). An empirical analysis of engineers' information behaviors. *Journal of the American Society for Information Science and Technology*, *61*(4), 640–658. http://doi.org/10.1002/asi.21290

.

Saúl, L.A., López-González, M.A., Moreno-Pulido, A., Compañ, V., Corbella, S., Feixas, G. (2012). Bibliometric review of the repertory grid technique: 1998–2007. *Journal of Constructivist Psychology*, *25*(2), 112-131.

Savolainen, R. (1993). The sense-making theory: Reviewing the interests of a user-centered approach to information seeking and use. *Information processing & management*, *29*(1), 13-28. doi:10.1016/0306-4573(93)90020-E

Savolainen, R. (2009). Small world and information grounds as contexts of information seeking and sharing. *Library & Information Science Research*, *31*(1), 38-45. doi:10.1016/j.lisr.2008.10.007

Scott, S. D., Graham, T. C., Wallace, J. R., Hancock, M., & Nacenta, M. (2015, February). Local remote collaboration: Applying remote group awareness techniques to co-located settings. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing* (pp. 319–324). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2685564

Seebach, C., Beck, R., & Pahlke, I. (2011). Situation awareness through social collaboration platforms in distributed work environments. *ICIS 2011 Proceedings*. Retrieved from http://aisel.aisnet.org/icis2011/proceedings/eastmeetswest/2

Serce, C., Alpaslan, F. N., Swigger, K., Brazile, R., Dafoulas, G., & Lopez, V. (2009). Dynamics of collaborative work in global software development environment. In: Cunningham, P., & Cunningham, M., (eds*.) eChallenges e-2009 Conference Proceedings* (paper 135*)*. Dublin: IIMC International Information Management Corporation Ltd.

Sharp, H., Giuffrida, R., & Melnik, G. (2012). Information flow within a dispersed agile team: A distributed cognition perspective. In C. Wohlin (Ed.), *Agile Processes in Software Engineering and Extreme Programming* (pp. 62–76). Berlin: Springer Berlin Heidelberg

Sharp, H., & Robinson, H. (2008). Collaboration and co-ordination in mature eXtreme programming teams. *International Journal of Human-Computer Studies*, *66*(7), 506–518.

Shen, Z., Lyytinen, K., & Yoo, Y. (2014). Time and information technology in teams: a review of empirical research and future research directions. *European Journal of Information Systems*. Retrieved from http://www.palgrave-journals.com/ejis/journal/vaop/ncurrent/abs/ejis20148a.html

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63–75.

Shrivastava, S.V., & Date, H. (2010). Distributed agile software development: A review. *arXiv preprint arXiv:1006.1955*. Retrieved from http://arxiv.org/abs/1006.1955

Siebdrat, F., Hoegl, M., & Ernst, H. (2014). Subjective distance and team collaboration in distributed teams. *Journal of Product Innovation Management*, *31*(4), 765–779.

Skelsey, G. (2014). Maslow’s hierarchy of needs - the sixth level. *Psychologist*, *27*(12), 982–983.

Sonnenwald, D. H. (2006). Challenges in sharing information effectively: examples from command and control [text]. Retrieved October 7, 2014, from http://www.informationr.net/ir/11-4/paper270.html

Sonnenwald, D. H. (1995). Contested collaboration: A descriptive model of intergroup communication in information system design. *Information Processing & Management*, *31*(6), 859–877.

Sonnenwald, D. H., Maglaughlin, K. L., & Whitton, M. C. (2004). Designing to support situation awareness across distances: an example from a scientific collaboratory. *Information Processing & Management*, *40*(6), 989-1011. doi:10.1016/j.ipm.2003.10.002

Sonnenwald, D. H., & Pierce, L. G. (2000). Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control. *Information Processing & Management*, *36*(3),461-479. doi:10.1016/S0306-4573(99)00039-4

Sonnenwald, D. H., Söderholm, H. M., Welch, G. F., Cairns, B. A., Manning, J. E., & Fuchs, H. (2014). Illuminating collaboration in emergency health care situations: Paramedic-physician collaboration and 3D telepresence technology. *Information Research*, *19*(2).

Spasser, M. A. (1999). Informing information science: The case for activity theory. *Journal of the American Society for Information Science*, *50*(12), 1136–1138. http://doi.org/10.1002/(SICI)1097-4571(1999)50:12<1136::AID-ASI17>3.0.CO;2-0

Staples, D. S., & Jarvenpaa, S. L. (2000, December). Using electronic media for information sharing activities: a replication and extension. In *Proceedings of the twenty first international conference on Information systems* (pp. 117–133). Association for Information Systems. Retrieved from http://dl.acm.org/citation.cfm?id=359729

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of experimental social psychology*, *31*(3), 244–265.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of personality and social psychology*, *48*(6), 1467-1478.

Stasser, G., & Titus, W. (2003). Hidden profiles: A brief history. *Psychological Inquiry*, *14*(3-4), 304–313.

Stewart, D. D., & Stasser, G. (1995). Expert role assignment and information sampling during collective recall and decision making. *Journal of personality and social psychology*, *69*(4), 619-628.

Stvilia, B., Hinnant, C. C., Wu, S., Worrall, A., Lee, D. J., Burnett, K., & Marty, P. F. (2015). Research project tasks, data, and perceptions of data quality in a condensed matter physics community. *Journal of the Association for Information Science and Technology*, *66*(2), 246–263. http://doi.org/10.1002/asi.23177

Stvilia, B., & Jörgensen, C. (2010). Member activities and quality of tags in a collection of historical photographs in Flickr. *Journal of the American Society for Information Science and Technology*, *61*(12), 2477–2489. http://doi.org/10.1002/asi.21432

Suchman, L. A. (1987). *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press.

Summers, J. K., Humphrey, S. E., & Ferris, G. R. (2012). Team member change, flux in coordination, and performance: Effects of strategic core roles, information transfer, and cognitive ability. *Academy of Management Journal, 55*(2), 314–338.

Sun, J. (2012). Why different people prefer different systems for different tasks: An activity perspective on technology adoption in a dynamic user environment. *Journal of the American Society for Information Science and Technology, 63*(1), 48–63. http://doi.org/10.1002/asi.21670

Suthers, D., Medina, R., Vatrapu, R., & Dwyer, N. (2007, July). Information sharing is incongruous with collaborative convergence: The case for interaction. In *Proceedings of the 8th international conference on Computer supported collaborative learning* (pp. 715–717). International Society of the Learning Sciences.

Swisher, K. (2013). "Physically together": Here's the internal Yahoo no-work-from-home memo for remote workers and maybe more. *AllThingsD, February, 22*. Retrieved from http://economics.uwo.ca/fubar_docs/jan_june13/physicallytogether_mar1.pdf

Szirtes, T. (2011). Managing knowledge sharing -- case study of a consulting firm. *Society and Economy, 34*(1), 139-160. doi:http://dx.doi.org/10.1556/SocEc.2011.0009

Tabak, E., & Willson, M. (2012). A non-linear model of information sharing practices in academic communities. *Library & Information Science Research, 34*(2), 110–116.

Talja, S. (2002). Information sharing in academic communities: Types and levels of collaboration in information seeking and use. *New Review of Information Behavior Research, 3(1)*, 143– 159.

Talja, S., & Hansen, P. (2006). Information sharing. In *New directions in human information behavior* (pp. 113–134). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/1-4020-3670-1_7.

Tang, J. C., Zhao, C., Cao, X., & Inkpen, K. (2011, March). Your time zone or mine?: A study of globally time zone-shifted collaboration. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 235–244). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1958860

Taylor, R. S. (1968). Question-negotiation and information seeking in libraries. *College & research libraries, 29*(3), 178-194.

Tenopir, C., & King, D. W. (2004). *Communication patterns of engineers*. John Wiley & Sons.

Tenopir, C., King, D. W., Spencer, J., & Wu, L. (2009). Variations in article seeking and reading patterns of academics: What makes a difference? *Library & Information Science Research*, *31*(3), 139–148.

The Standish Group. (2013). Chaos Manifesto 2013. Retrieved September 30, 2015 from https://larlet.fr/static/david/stream/ChaosManifesto2013.pdf

Toomela, A. (2000). Activity theory is a dead end for cultural-historical psychology. *Culture & Psychology*, *6*(3), 353-364.

Toomela, A. (2008). Commentary: activity theory is a dead end for methodological thinking in cultural psychology too. *Culture & Psychology*, *14*(3), 289-303.

Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of small-group development revisited. *Group & Organization Management*, *2*(4), 419–427.

Van der Veer, R., & Yasnitsky, A. (2011). Vygotsky in English: What still needs to be done. *Integrative Psychological and Behavioral Science*, *45*(4), 475–493.

Van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, *58*, 515–541.

Verner, J. M., Overmyer, S. P., & McCain, K. W. (1999). In the 25 years since The Mythical Man-Month what have we learned about project management? *Information and Software Technology*, *41*(14), 1021–1026.

Voida, A., Bos, N., Olson, J., Olson, G., & Dunning, L. (2012, May). Cross-cutting faultlines of location and shared identity in the intergroup cooperation of partially distributed groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3101–3110). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2208724

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. (Cole, M., John-Steiner, V., Scribner, S., & Souberman, E., eds.). *Cambridge, MA: Harvard University Press.*

Wagstrom, P., & Datta, S. (2014, May). Does latitude hurt while longitude kills? geographical and temporal separation in a large scale software development project. In *Proceedings of the 36th International Conference on Software Engineering* (pp. 199–210). New York, NY, USA: ACM. http://doi.org/10.1145/2568225.2568279

Wanberg, J., Javernick-Will, A., Chinowsky, P., & Taylor, J. E. (2015). Spanning cultural and geographic barriers with knowledge pipelines in multinational communities of practice. *Journal of Construction Engineering and Management*, *141*(4), 04014091:1-10. http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000955

Wang, C. Y., Yang, H. Y., & Chou, S. T. (2008). Using peer-to-peer technology for knowledge sharing in communities of practices. *Decision Support Systems*, *45*(3), 528–540.

Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, *20*(2), 115–131. http://doi.org/10.1016/j.hrmr.2009.10.001

Weedman, J. (1993). On the "isolation" of humanists: A report of an invisible college. *Communication Research*, *20*(6), 749–776. http://doi.org/10.1177/009365093020006001

Wertsch, J. V., Del Rio, P., & Alvarez, A. (1995). *Sociocultural studies of mind*. Cambridge University Press.

Widén, G., & Hansen, P. (2012). Managing collaborative information sharing: bridging research on information culture and collaborative information behaviour. *Information research*, *17*(4).

Widén-Wulff, G., & Davenport, E. (2007). Activity systems, information sharing and the development of organizational knowledge in two Finnish firms: an exploratory study using activity theory. *Information research*, *12*(3).

Wilson, T.D. (1999). Models in information behaviour research. *Journal of documentation*, *55*(3), 249 - 270. doi:10.1108/EUM0000000007145

Wilson, T. D. (2006). A re-examination of information seeking behaviour in the context of activity theory. *Information Research*, *11*(4).

Wilson, T.D. (2007) Editorial. *Information Research*, 12(3), editorial E123 [Available at: http://informationr.net/ir/12-3/editor123.html]

Wilson, T. D. (2008). Activity theory and information seeking. *Annual Review of Information Science and Technology*, *42*(1), 119-161. doi:10.1002/aris.2008.1440420111

Wilson, T. D. (2010). Information Sharing: An Exploration of the Literature and Some Propositions. *Information Research: An International Electronic Journal*, *15*(4), n4.

Wilson, T. D. (2013). (ed.) *Theory in information behaviour research*. Sheffield, UK: Eiconics LTD.

Witherspoon, C. L., Bergner, J., Cockrell, C., & Stone, D. N. (2013). Antecedents of organizational knowledge sharing: A meta-analysis and critique. *Journal of Knowledge Management*, *17*(2), 250-277. doi:http://dx.doi.org/10.1108/13673271311315204

Xu, Y. (2007). The dynamics of interactive information retrieval behavior, Part I: An activity theory perspective. *Journal of the American Society for Information Science and Technology*, *58*(7), 958–970. http://doi.org/10.1002/asi.20569

Xu, Y., & Liu, C. (2007). The dynamics of interactive information retrieval, Part II: An empirical study from the activity theory perspective. *Journal of the American Society for Information Science and Technology, 58*(7), 987–998. http://doi.org/10.1002/asi.20574

yahia, N. B., Bellamine, N., & Ghézala, H. B. (2012). On the Convergence of Collaboration and Knowledge Management. *arXiv:1202.6104 [cs]*. Retrieved from http://arxiv.org/abs/1202.6104

Yamagata-Lynch, L. C. (2010). Understanding cultural historical activity theory. In *Activity Systems Analysis Methods* (pp. 13–26). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4419-6321-5_2

Yang, T.-M., & Maxwell, T. A. (2011). Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors. *Government Information Quarterly*, *28*(2), 164–175.

Zahedi, M., & Babar, M. A. (2014, May). Knowledge sharing for common understanding of technical specifications through artifactual culture. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering* (p. 11). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2601293

Zaitsev, A., Gal, U., & Tan, B. (2014, December). Boundary objects and change in agile projects. *Proceedings of the 25th Australasian Conference of Information Systems ACIS 2014*, Auckland, New Zealand: ACIS. Retrieved from http://aut.researchgateway.ac.nz/handle/10292/8043

Zins, C. (2006). Redefining information science: from "information science" to "knowledge science." *Journal of Documentation*, *62*(4), 447–461.

Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, *58*(4), 479-493. doi:10.1002/asi.20508

Zuccala, A. (2006). Modeling the invisible college. *Journal of the American Society for Information Science and Technology*, *57*(2), 152–168.

Appendices

APPENDIX A: INTERVIEW PROTOCOL FOR PILOT ONE

STUDY INSTRUMENT FOR PILOT ONE
Pilot Interview & Notes
Interviewee:
Date and time:
Call identifier:

Section A: general questions about the overall activity (Mwanza's Eight Step Model to guide Systems Design, 2001):

1. *"Activity* of interest: (project)
   What sort of activity is it?
2. *Objective* of activity
   Why is this activity taking place?
3. *Subjects* in this activity
   Who is involved in carrying out this activity?
4. *Tools* mediating the activity
   - By what means are the subjects carrying out this activity?
   - What Tools do the Subjects use to achieve their Objective and how?
5. *Rules and regulations* mediating the activity
   - Are there any cultural norms, rules, or regulations governing the performance of this activity?
   - What Rules affect the way the Subjects achieve the Objective and how?
6. *Division of labor* mediating the activity
   - Who is responsible for what when carrying out this activity and how are the roles organized?
   - How does the Division of Labor influence the way the Subjects satisfy their Objective?
7. *Community* in which activity is conducted
   - What is the environment in which this activity is carried out?
   - How do the Tools in use affect the way the Community achieves the Objective?
   - What Rules affect the way the Community satisfies their Objective and how?
   - How does the Division of Labor affect the way the Community achieves the Objective?
8. What is the desired *Outcome* from carrying out this activity?" (p. 6)

Section B.: questions about information sharing – (Activity Checklist from Kaptelinin et al., 1999):

“Means and ends - the extent to which information sharing facilitates and constrains the attainment of users’ goals and the impact of information sharing on provoking or resolving conflicts between different goals

Social and physical aspects of the environment – integration of target technology with requirements, tools, resources, and social rules of the environment.

Learning, cognition, and articulation – internal versus external components of activity and support of their mutual transformations by information sharing

Development – developmental transformation of the foregoing components as a whole.” (p.33)

1. When did you last share any project information with a team member?
2. What was your intent when you shared <document x> with <colleague a>?
3. How did this fit into the project work?
4. Why do people from your project share information with each other? Why do you?
5. How often does this occur during the course of a project? Do you usually share information with others on the project?
    a. Why? And under what circumstances?
    b. With everyone or just some of the people?
6. And do they (or others) share information with you?
7. Under what circumstances?
    a. (is it symmetrical - do pairs of people share in both
    b. Directions or is it asymmetrical)
8. For analysis in this PhD research, are you willing to share with me any documents, emails or any other specific examples of information:
    a. that you have shared with other project team members, or
    b. That other project team members have shared with you?
9. Please tell me about the experience of sharing information on this project.
    a. (e.g., is it positive, helpful, problematic - interviewer will probe if needed)

<u>Section C: Information about the interviewee and demographics</u>

1. What is your educational background?
2. Tell me about your professional experiences within the company
3. Information about you and the project , e.g.,
    - How long have you been on the project?
    - What is your role on the project?
4. How do you typically communicate with others on the project?
5. Any other thoughts that come to mind?

APPENDIX B: INTERVIEW PROTOCOL FOR PILOT TWO

Interview Protocol: Information Sharing in Virtual Collaboration First Interview

1. Introduction

The purpose of my PhD research project is to explore and investigate information sharing among members of a virtual software development team. As part of my project I am gathering data about information sharing. I am interested in your experiences and perspectives about information as you have worked in a software engineering project. I will be recording audio of our conversation and taking notes so that I have a record of what we talk about, and perform subsequent analysis. All comments and responses will be treated confidentially. The names of individual persons are not required in any of the responses, although I am interested in what role the person plays, and things like their organizational relationship to you. I will be using a screen sharing session, which will also be recorded, to discuss these ideas.

As a reminder, your participation in this project is voluntary. If you do agree to participate, you can withdraw from participation at any time during the project without comment or penalty. Your decision to participate or to withdraw will in no way impact upon your current or future relationship with QUT or with this company.

Do you have any questions about your participation in my study?

Our conversation is a structured interview, which means I have a framework to guide our discussion and move toward as much precision as possible. You can choose the level of detail you would like to discuss. Through this, I will be trying to understand you and your perspective in your own terms, not to collect the "right" answers. I will be asking you to make a series of systematic comparisons in order to explore this topic.

Do you have any (other) questions before we start?

3. Elicitation of Elements (~10)

Please think about your project and what information you share, as well as the information others share with you.

Can you think of examples of?

Shared information in your project?
What would you call that?
We are going to put each one of these on a "card"

Prompts if needed:

Information that you shared with one person?
Information that you shared with multiple people?
Information that someone else shared with you?
Information that you didn't share?
Information that you requested be shared?
Information that you requested be shared but didn't receive?
Information you received that was not useful to you?
Information you received that you wish you didn't?
Critical information?
Information that came too late?
Information that you didn't understand?
Information that you already knew about?

3. Elicitation of Constructs (~ 7-12)

Now, I would like to explore something about the characteristics of these types of information you named, and what makes some of them similar to each other, or different from each other. I would like to know how you think about them.

We are now going to look at three of the items at a time that you identified. Which two of these are the *same* in some way, and *different* from the third?

What is it that item #1 and #2 have in common? (as opposed to #3)

What makes item #3 different?

Now let's look at another trio of items… < continue until the interviewee can do no more >

4. Wrap-up and Conclusion

Thank you very much for your time and for sharing your insights and experiences with me. As a reminder, all or you comments and responses will be treated confidentially and will be aggregated with the responses of others for analysis.

I will send you the transcript of our conversation with you, so that you can make any corrections or add additional comments for clarification. I will also contact you in a few weeks with a short follow-up survey.

Do you have any questions?

Thanks again.

Second Interview

The purpose of the second interview is to have the participant provide ratings for the single repertory grid -- elements (the nouns) and constructs (the verbs). It is a consolidation created by me of all of the interviews that were conducted with this team. I will be asking you to provide a rating from 1 (the characteristic at the left pole) to 5 (the characteristic at the right pole) for each element. We will go one at a time.

Do you have any questions?

5. Rating of Elements on Constructs:
1 2 3 4 5
6. Wrap-up and Conclusion
Thank you very much for your time and for sharing your insights and experiences with me. As a reminder, all of your comments and responses will be treated confidentially and will be aggregated with the responses of others for analysis.

I will send you the matrix that you just completed, so that you can make any corrections or add additional comments for clarification. Do you have any questions?

Thanks again.

APPENDIX C: INTERVIEW PROTOCOL FOR MAIN STUDY

Interview Protocol: Information Sharing in Virtual Collaboration
11/2013 Laura Anderson

Initial Interview

1. Introduction

The purpose of my PhD research project is to explore and investigate information sharing and collaboration among members of a virtual software development team. As part of my project I am gathering data about information sharing. I am interested in your experiences and perspectives about information as you have worked in a software engineering project. I will be recording audio of our conversation and taking notes so that I have a record of what we talk about, and perform subsequent analysis. All comments and responses will be treated confidentially. The names of individual persons are not required in any of the responses, although I am interested in what role the person plays, and things like their organizational relationship to you. I will be using a screen sharing session, which will also be recorded, to discuss these ideas.

As a reminder, your participation in this project is voluntary. If you do agree to participate, you can withdraw from participation at any time during the project without comment or penalty. Your decision to participate or to withdraw will in no way impact upon your current or future relationship with QUT or with this company.

Do you have any questions about your participation in my study?

Our conversation is a structured interview, which means I have a framework to guide our discussion and move toward as much precision as possible. You can choose the level of detail you would like to discuss. Through this, I will be trying to understand you and your perspective in your own terms, not to collect the "right" answers. I will be asking you to make a series of systematic comparisons in order to explore this topic.

Do you have any (other) questions before we start?

2. Elicitation of Elements (~10)

Please think about your project.

Can you think of examples of?

What is shared in your project?
What would you call that?
We are going to put each one of these on an index "card"

Prompts if needed:

Information that you shared with one person?

Information that you shared with multiple people?
Information that someone else shared with you?
Information that you didn't share?
Information that you requested be shared?
Information that you requested be shared but didn't receive?
Information you received that was not useful to you?
Information you received that you wish you didn't?
Critical information?
Information that came too late?
Information that you did not understand?
Information that you already knew about?

3. Elicitation of Constructs (~ 7-12)

Now, I would like to explore something about the characteristics of these types of information you named, and what makes some of them similar to each other, or different from each other. I would like to know how you think about them.

We are now going to look at three of the items at a time that you identified. Which two of these are the *same* in some way, and *different* from the third?

What is it that item #1 and #2 have in common? (as opposed to #3)

What makes item #3 different?

Now let us look at another trio of items…

4. Wrap-up and Conclusion

Thank you very much for your time and for sharing your insights and experiences with me.

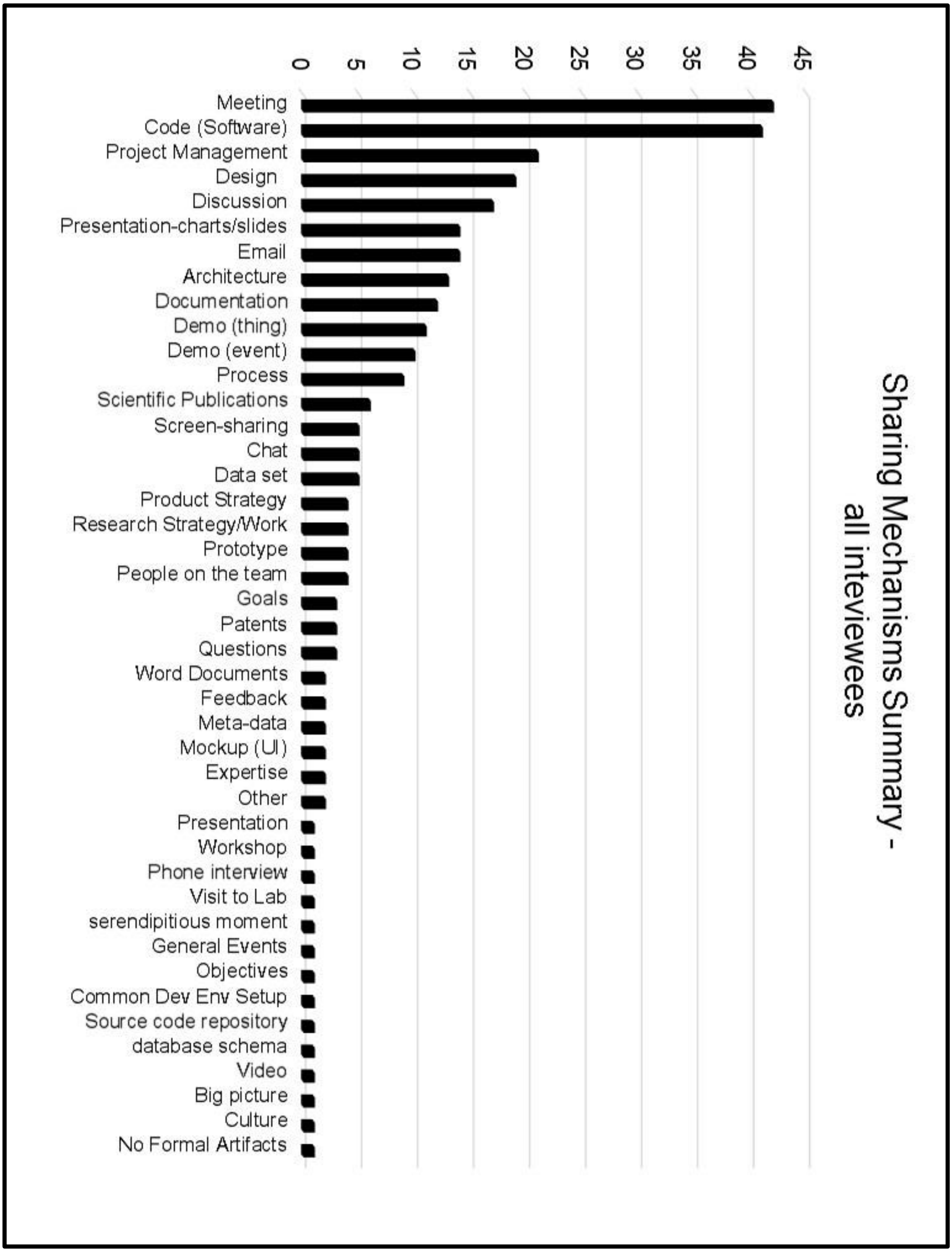
I am interested in examples of artifacts (or databases) that are shared in the project, and would be appreciative if you could send any of those to me. I would, of course, treat these very confidential and use them only for this study.

As a reminder, all of your comments and responses will be treated confidentially and will be aggregated with the responses of others for analysis. I will send you the transcript of our conversation with you, so that you can make any corrections or add additional comments for clarification. I will also contact you in a few weeks with a short follow-up survey.

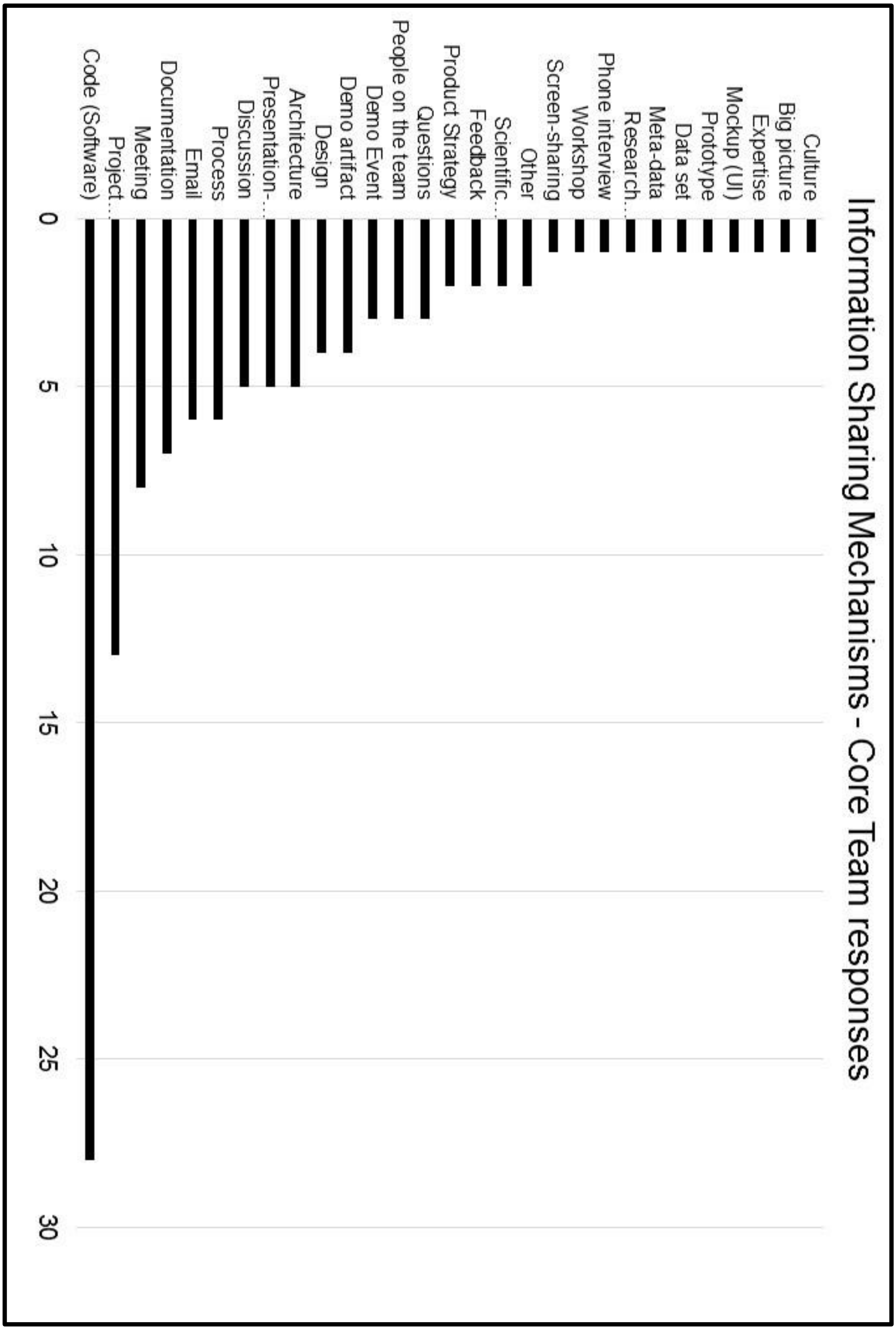
Do you have any questions?

Thanks again.

APPENDIX D: INFORMATION SHARING MECHANISMS SUMMARY – ALL PARTICIPANTS



APPENDIX E: INFORMATION SHARING MECHANISMS SUMMARY – CORE TEAM



APPENDIX F: INFORMATION SHARING MECHANISMS SUMMARY – EXTENDED STAKEHOLDERS