Studying Technologically Dense Environments through Documenting Practices

Carsten Østerlund
Syracuse University
costerlu@syr.edu

Steve Sawyer
Syracuse University
ssawyer@syr.edu

Elizabeth Kaziunas
Syracuse University
eskaziun@syr.edu

Abstract:

This paper explores the methodological aspects of studying distributed work in technology dense environments by focusing on the more tangible aspects of this work, namely documents. The authors outline a process oriented theory of documenting work by drawing on document scholarship, workplace studies, communication genres, and science and technology studies. Using data from a pilot study of virtual organizing among social scientists, we outline a document-centric methodology. We show how 1) researchers should engage in an initial mapping of documents before starting to track them; 2) the ongoing flow of virtual organizing only becomes apparent by triangulating the digital flow of documents, observation of tangible documents (e.g., paper) and repeated behavioral inquires; 3), documents supporting virtual organizing do not serve as stable information artifacts, but rather become snapshots in time, part of the general flow of work across numerous documents and applications. We evaluate how best to combine document tracking with interviews and participant observation, and discuss the challenges and benefits associated with digital instrumentation, practicality, privacy, verifiability and reliability.

Introduction[1]

Distributed work in technologically dense environments is hard to observe. A focus on the more tangible aspects of this work, namely documents, provides a useful lens into the work practices of organizational members in general, and those working in heterogeneous networks in particular. We define documents as typified and material communication, whether electronic, paper-based, wall mounted or set in stone, invoked in response to recurrent situations (Østerlund 2008; Østerlund and Boland 2009). Studying documents in work allows us to position people's immediate activities and situated routines in their larger social and organizational context. As documents carry institutional structures and point to both past and future activities they open a window to larger organizational practices.

---

[1] In part the present paper builds on earlier work presented at the 26th EGOS Colloquium 2010, Lisbon, Sub-Theme 18.

The notion of document serves as a lens into the socio-technical or socio-material nature of what organizational members do, day in and day out. Documents are socio-technical in that they are both material – and embody the technical infrastructure – and social – as they embody both the work practices and shared understanding of those involved. For example, our production and circulation of the document in front of you involved the technology of word processors, several different computers, Google documents, hard copies, email messages and a variety of file formats. We even touched a book in the process.

Your reading of this paper involves numerous other technologies; you are likely reading either a printed or digital version of the paper that you received in your email inbox. Some of you might have accessed this paper's file through the EGOS online system. You may even be struggling to read this on the limited screen of your mobile phone. Shared social practices are reflected in the degree to which you, the reader, and we, the authors, understand and share common knowledge about the form and contents of the genre of paper submissions in general and EGOS papers in particular and reflect this knowledge in this document. The work we have done, and that you are doing, represents the basics of our coordination in action. And, the various material forms of the submission represent some of the technology dense environment supporting the EGOS conference and the broader organization studies field.

The production and consumption of this paper involves both the work of documenting and document work. The work of documenting falls close to the definition of the verb, to "document," describing the act of providing factual or substantive support for statements made or hypotheses proposed; or to equip with exact references to authoritative supporting information (Merriam-Webster Online Dictionary). At the same time we engaged in document work involving the production, use, collection, classification, storage, retrieving and dissemination of documents within and across organizational settings. Documents are not solely accounts of practice. They are also accounting for practice. Documents offer models for practices and thus allow people to coordinate their distributed work.

If we accept that documenting work opens a conceptual door to work practices in technologically dense environments the methodological question immediately arise: How do we best study documenting work? The present paper attempts to address this broader methodological question by discussing the theoretical framework, data gathering and analytical techniques developed for an ongoing study of distributed social scientists. The paper focuses on the *methodological* benefits and challenges of studying social scientists' coordination through a documenting work lens.

Theoretical background

Documents provide a window into the distributed coordination of scientists and the digital infrastructure that supports them. The efforts to develop, share, use and store documents have been a focus of scholarly attention since the late 19th century. Much of this work focused on scholarly interaction and science practice. For example, more than a century ago many American and European scholars worked together to create environments and tools for international collaboration. During this period a large number of international scientific

associations and journals were founded, in part to increase sharing of knowledge. The increased proliferation of scientific collaboration and production created a need for tools to locate colleagues' work, to find publications, create and share collections of data, and coordinate collaborations. The notion of *document* emerged as the organizing concept for this new science and lead to a number of productive questions associated with the material manifestations, temporal and spatial production, distribution, inventory, statistics, preservation and use of documents (Otlet 1903; Otlet 1907).

More than 100 years on we find the same questions even more pressing. The emphasis has shifted from the printed text to the internet, seemingly with little attention to the useful insights from this path-breaking earlier work. Challenges associated with digital documents in their many manifestations, such as their production, distribution, inventory, preservation and uses, reflect attempts to facilitate highly distributed scientific collaborations. These challenges both replicate and magnify the issues first touched on at the beginning of the last century.

Distributed work is dynamic. Scientists engage with many documents in their daily activities, from research reports to email from colleagues to facilitate different types of virtual collaboration with different participants, at distinct times and places, and around different kinds of content. Several bodies of literature are relevant to this area including document scholarship (Brown and Duguid 1994; Buckland 2007; Frohmann 2007; Frohmann 2004; Harper 1998; Lund 2009; Østerlund and Boland 2009; Zacklad 2006), sociomateriality (Barad 2003; Orlikowski 2007; Suchman 2007), workplace studies (Forsythe 1999; Pollock, Williams, and . 2010; Randall, Harper, and Rouncefield 2007; Suchman 1987; Suchman 2000; Suchman 2007), science and technology studies (Hine 2006; Knorr Cetina 1999; Latour 1987; Latour and Woolgar 1986); communication genres (Bazerman 1995; Crowston and Kwasnik 2003; Orlikowski and Yates 1994; Østerlund 2007; Swales 1990; Yates and Orlikowski 2007), and boundary objects (Ackerman and Halverson 2004; Carlile 1997; Carlile 2004; Star and Griesemer 1989).

Not all these research strands draw on the notion of documents. What brings them together is an emphasis on the *practices* associated with communicative artifacts and their *temporal*, *spatial* and *material* dimensions. Only through studies of people's unfolding activities can we understand how documents are produced and reproduced. Derived from such a practice perspective these scholars insist that documents are localized and contingent. As they move from place to place, get used at different times and transform from one material form to another their meaning, purpose and uses alter.

The document scholarship literature develops out of the library and information science community with roots back to work by Otlet (1903; Otlet 1907), Briet (1951; Briet, Day, Martinet, and Anghelescu 2006) and other early scholars of documentation. The concept faded with the advent of the computer and has only recently been revised by a small but growing number of scholars. This recent debate has predominantly approached documents conceptually and rarely engaged in empirical studies. For instance, Frohmann define documents as "different material kinds of temporally and spatially situated bundles of

inscriptions embedded in specific kinds of cultural practices" (Frohmann 2004: 137). Drawing on Foucault and social scientists, such as Bourdieu and Garfinkel, he articulates a document perspective which emphasizes the localized, contingent, and difficult labor involved in producing stable phenomena that can be extended across contexts. It focuses on the work involved in aligning the many heterogeneous elements of a community or setting to stabilize a phenomenon. Many document scholars place similarly emphasis on the material aspects of documents. The French scholar Zacklad, for instance, writes that a document is "a semiotic product transcribed or recorded on a perennial substrate, which is endowed with specific attributes intended to facilitate the practices associated with its subsequent utilization in the framework of the distributed communicational transactions" (Zacklad 2006: 217).

Recent genre theories provide a comparable emphasis on practice when studying communication artifacts. Drawing on pragmatic and practice-oriented approaches, a number of researchers in cultural, communication, and system design studies have begun approaching genres as classes of communicative practices (Bahtia 1993; Bakhtin 1986; Bakhtin 1996; Bazerman 1995; Miller 1984; Swales 1990). This literature has not paid as much attention to the material aspects of documents. In fact they do not restrict such communicative practices to material artifacts, or in Zacklad's words, "the perennial substrate." Meetings and other habitual oral communication can be regarded as communication genres. Instead, this literature offers a process perspective which points to how communicative practices are organized sequentially to facilitate evolving practices and multiple parties in a community. In Bazerman's words: "Only a limited range of genres may appropriately follow upon another in particular settings, because the success conditions of the actions of each require various states of affairs to exist" (Bazerman 1995: 98). Documents thus do not stand alone. Their temporal and spatial enactment and combination with other documents matter for how they support the activities of specific groups (Østerlund 2007).

Numerous scholars in science and technology studies contribute to this debate through rich empirical studies. Latour and Woolgar's classic study of laboratory life (Latour and Woolgar 1986) nicely illustrates the power of focusing on the practices associated with the production, translation and circulation of documents with an emphasis on their temporal, spatial and material dimensions. They describe the scientist's desk as the hub of a larger productive system, the laboratory. Latour and Woolgar pay attention to what documents can be found on those desks, where the documents originate, and how scientists juxtapose documents. In the laboratory's production of scientific knowledge these documents are not stable entities but are constantly manipulated and transformed. Data are gathered from inscription devices and then translated from one document to another. In the process some things are deleted, others isolated, detected or named. To Latour and Woolgar it becomes important to track those documents over time, place and material manifestations to understand the production of knowledge as documents transform, combine and get reshuffled.

The notion of provenance contributes to such a process-oriented and practice-based approach to documents. Historically, the concept of provenance belonged to archival research and refers to where something comes from, as in where did a document, a bit of data, or artifact first appear, but also to a kind of genealogy of artifacts (e.g., artwork), documents, and ideas

(Baudoin 2008; Lonsdale, Jensen, Wynn, and Dedual 2010; Sweeney 2008). More recently the concept has been adopted by the information management community and computer science (Blanc-Brude and Scapin 2007; Lynch 2001; Shen, Fitzhenry, and Dietterich 2009). An "original" text may be authored by the user or assemblage from multiple sources. Often content flows from application to application and document to document, constantly recycled, reworked and repackaged. In computer science the notion of provenance is used in studies of document management and retrieval but also as a way to determine privacy and document ownership. Supporting, tracking, and visualizing these relations among documents are essential to providing better support for information workers.

Documenting practices: Situating ourselves in this broader debate we can regard documents as material manifestations of individual, group, or communal practices. In many professions such as academia documents even make up the main outcome of a community's practices. It is through repeated documenting practices participants stabilize the relations among heterogeneous elements in their community. In other words, documenting practices become a physical expression of the very social relations that define their field of practice. Those relations are constantly produced and reproduced through the ongoing documenting practices of communal participants.

Following this line of thinking one can ask, what do these material relations do? How do documents work as material manifestations of social relations? Taking inspiration from John Law's concept of 'relational materiality' the answer becomes (Law 2004): Documents enact presence, manifest absence, and Otherness. Presence refers to whatever is present in a document such as descriptions or data relevant for a community. Absence points to whatever is absent but manifest in the document. It can be the activities, people, or things to which the descriptions or data in the document refers. Otherness include an endless number of processes and contexts necessary to whatever is present and absent in the document but also invisible to it. It may range from things that everyone engaged with the documents knows through experience and thus do not need to mention or matters that are actively suppressed in order to realize specific practices.

Together the presence, absence, and Otherness captured by a document constitute models of and models for practices. Documents offer representations of the world by making part of that realty present and absent in the text, while leaving many other things outside the field of attention. By the same stroke this makes them representations for practices as it calls people's attention to what matters, what should be acted upon, and what can be left untouched.

As material manifestations of relations, documents are not stable artifacts but constitute more or less precarious chains of documenting practices. Some things get deleted, pushed into invisibility, as new things, perspectives, or activities take precedence. Sometimes these relations take shape within the confines of one document as it evolves over time. The production of this conference paper has gone through a number of iterations each making various activities of the authors' research and thinking present, absent or simply pushed to the wayside (i.e., Otherness). Equally often, such evolving relations can be traced through the provenance of a document. The present paper draws on a number of transcribed interviews,

field notes, analytical memos, grant proposals, email exchanges, and published papers. One can track the practices and coordination among the authors through these numerous documents as content from one document gets translated or copied to another, repositioned, and builds into other configurations.

The content of a document is not the only marker of its manifest relations. Their temporal, spatial, and material position also makes a difference in terms of what relations a documents call attention to and what is left unexplored. A document in a trash bin does not serve the same role in a collaboration as one found on the author's easy chair. The relations among a group of participants articulated by a document may be negated or simply forgotten as the text makes its way into the trash bin. Less radical, the way participants choose to classify and file a document among other texts has significance to how it works as a model for and of a community's collaboration and coordination. In the same way, the manifest relations of a document can change over time as it gets transformed from one materiality to another. Converting handwritten notes on a pad of paper into an electronic document changes its potential relations. The content may remain the same but its position in a community and the relations it makes present, absent, or neglects all together are likely to alter. From being a material manifestation of an individual's work it may become something that a whole group can draw on and tie into their collective practices.

Consequently, a focus on documenting practices shifts our attention away from the concepts of documents as stable artifacts and replaces it with a process perspective. In this view, documents change with unfolding practices or they become moments or snapshots in time, part of the general flow of work across numerous documents and applications. Such a conceptualization of documenting practices makes it relevant for studies of people's working and organizing in technology dense environments. As documents call attention to various parts of the world and produce representations of and for practices they allow collaborators to synch their attention and coordinate their activities. By studying the practices associated with document production, classification, storage, retrieval, and use we can develop an understanding of the dynamic and shifting relations that facilitate collaboration and coordination.

Method

So, how *do* we best study documenting work? The answer may seem tantalizingly straight forward. You gather a pile of what social scientists drop left and right and start digging through it. But if you step back and begin scanning through your qualitative method books you will realize that documents subsist as a lower caste in field research. Most chapters and articles will help the reader refine their interview and participant observation skills. Documents and other artifacts are often addressed in passing under headlines such as "secondary sources" or "unobtrusive techniques," if at all. Many an ethnographer has more or less haphazardly gathered piles and piles of documents in the field, whether they did single-sited, multi-sited, participant generated or historical ethnography (Geiger and Ribes 2011). Yet, these documents often get collected without a good understanding of the larger body of documents from which they are drawn or the technical infrastructure supporting them.

Consequently, new integrative strategies have to be developed to track work practices in distributed and technologically dense environments.

Pilot Study: Distributed work and organizing among social scientists

In an effort to develop the methodological tools suited for a study of documenting work we are engaging in a two phase pilot study. First, we have investigated our own distributed practices. The subjects include the three coauthors (Carsten, Liz and Steve), our colleague (Caroline Haythornthwaite) from another major research university in the US who was in the UK for a visiting professorship, another colleague in Canada, and grant administrators at our respective institutions. Second, we are in the process of selecting four more collaborations among distributed social scientists. As part of this process we have to date interviewed ten scholars each engaged in two to three distributed collaborations on average, and participated in the meetings of one larger project. These scholars are located in the US, UK, Australia, and Scandinavia. Their distributed collaborations range from small groups focused on grant and article writing projects and a few larger collaborations with ongoing funding.

While it might be controversial to use ourselves and social science colleagues as research subjects we have two motives for doing so. First, studying people's documenting work potentially intrudes deeply into places where personal information dwells side by side with project relevant documents. Many people engage in work related activities in public and private spaces. Navigating this treacherous landscape requires some practice before exposing others to the instruments and efforts. Studying our own coordination in action enable us to try out different strategies, monitor their effects on the subjects and adjust our techniques as we learn the strengths and weaknesses of each approach.

Second, this pilot study is being pursued as part of a larger effort to better understand e-science (i.e., cyberinfrastructure) issues among social science scholars. Compared to the "big sciences" (e.g., physics, biology, chemistry) social science coordination has received little attention. Moreover, the latter group tends to be characterized by relative short term coordination and cooperation within and across groups with little permanent infrastructure support. In contrast, the virtual collaboration and information sharing in 'big science' collaborations has received significant attention in recent years including the building of large shared information repositories and joint access to instruments (e.g., telescopes, particle accelerators). Exploring the dynamics of our own local and virtual coordination offers a valuable stepping stone in designing and articulating the theoretical and methodological implications of the approach.

Data collection

The approach to data collection builds on our basic premise detailed above: documents are both central elements of organizational infrastructure (i.e., the structure of the field) and a part of practice -- the unfolding of coordination over time. Where a more traditional conception of infrastructure – as an information system -- tends to present documents as a static substrate: something upon which something else runs or operates (e.g., a railway), we see documenting work and the infrastructures they form as emerging from people's enacted practice, connected to activities and structures (Star and Ruhdeler 1996: 112). Nevertheless,

in the process of gathering data it can be helpful for a moment to think of the structure of the field as distinct from the process of practice. To understand the process of documenting work, it is important to have a sense of the overarching structure of the field in which these processes play out. The methodological consequence of this contextual embeddedness is a two-step research process: first, map the field of documents and second, follow the flow of their production and use. Figure 1 summarizes our document-centric data gathering strategy. The two light gray boxes belong to the mapping phase. The four darker boxes point to the tracking phase. We will discuss these two strategies in turn.
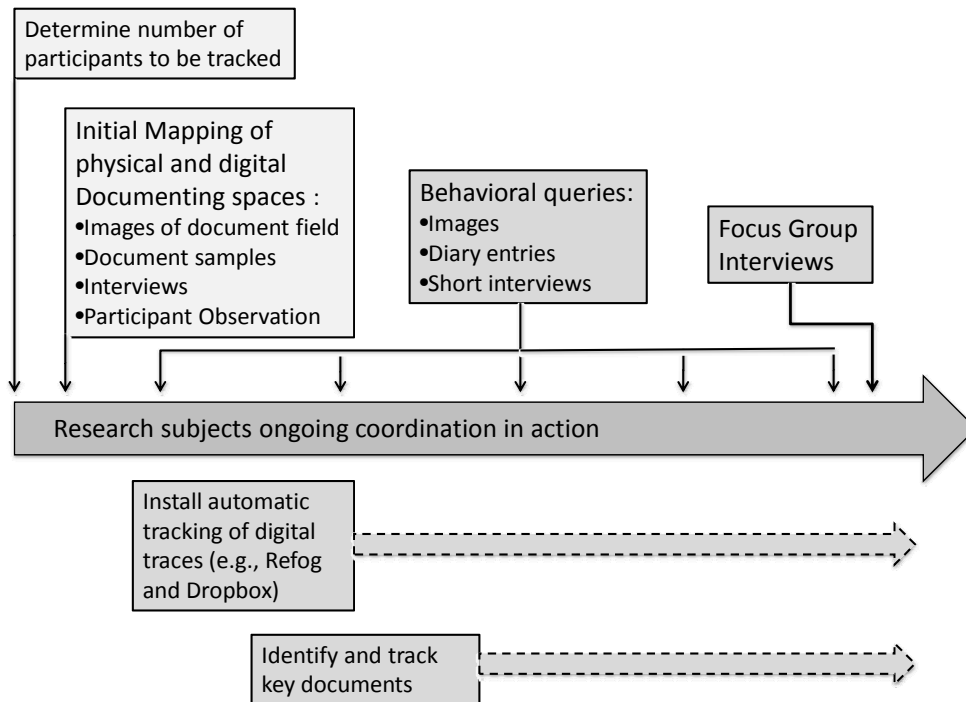
Determine number of participants to be tracked

Initial Mapping of physical and digital Documenting spaces :
•Images of document field
•Document samples
•Interviews
•Participant Observation

Behavioral queries:
•Images
•Diary entries
•Short interviews

Focus Group Interviews

Research subjects ongoing coordination in action

Install automatic tracking of digital traces (e.g., Refog and Dropbox)

Identify and track key documents

Figure 1: Document-centric data gathering strategy

Mapping the field of documents

The mapping involves all participants in a distributed collaboration. After we have determined the scope of the community, we pursue for each participant a combination of interviews and participant observation to detail document arrangements across both physical and digital spaces. The mapping is framed by the overarching question: Where do your documents live? The informant is asked to give a "tour" of their physical and electronic document spaces, which may take one from unwieldy piles of papers stacked on an office floor and stacks of journals along a window ledge to neatly structured desktop file folders and bulky email inboxes. It can be a good idea to start broad by asking to the participants' general documenting practices and gradually focus the questions to project specific document work.

We gather data about how the scholars' organize their desktops, home and work offices and filing system. Photos and videos serve as helpful tools. Then, we detail the organization of project-related documents, which spans physical and electronic spaces including shared repositories on organizational networks on in the cloud. In lieu of video, screen shots capture the organizational scheme of desktop project folders and subfolders as well as the

organization of email and shared repositories (e.g., MS Sharepoint, learning management systems, Google doc).

Interviews and observations can be conducted face-to-face or virtually. In the former case the researcher will record digital still images and video during the interview and observation. A digital camcorder captures both audio, physical documents, and their context, whether home, office or public space. The act of filming helps participants focus on and articulate the details of their document infrastructure and how it comes to life in their daily documenting work. When conducting the interview and observation virtually we find it beneficial to ask the informants to furnish the researcher with images of their field of documents before the interview, physical and electronic. These pictures provide a good starting point for the interview but will typically require the informant to supplement with additional pictures prompted by interview questions. Skype offers a convenient way to conduct online interviews and one finds a host of different applications allowing the researcher to record voice and images. A number of web conference systems offer the same capabilities (e.g., WebEx, Adobe Connect Pro).

Content analysis allows us to map the basic structure of the field and calibrate our data gathering of the scholars' ongoing collaboration. Only based on a detailed understanding of each participant's document infrastructure can we set up a comprehensive system to track document flow, and thus coordination in action. Let us illustrate the type of data the mapping can produce with a few examples from our pilot study.

The location of documents: In the interviews all participants distinguished multiple locations for our documenting work: the office, home office, home desks used for work, and public spaces (e.g., café or library). Among the three faculty members the office stands out as a dumping ground for documents, dead or alive. Steve leaves all documents in his office and only "checks out" items that he plans to work on over the next couple of days. Caroline and Carsten use their office for predominantly teaching and service related texts and inactive project documents. "It's a place where documents come to die." The home office holds more active document piles in Caroline, Liz and Carsten's case whereas Steve keeps his home empty of work documents with the exception of files he is actively work on at the moment.

A similar pattern emerges from our mapping of other social scientists. Most use their office as a main document repository. Some researchers use home for serious writing, while others struggle to keep it a work-free zone. However, no participants seem able to keep email, grading and some article writing out of their homes, not even the Scandinavians with strong cultural norms of guarding non-work time.

Piles and folders: Across these locations we all organize our documents in physical piles and electronic folders. The extent of their organization and use vary among each member. One faculty member, Caroline keeps administrative and teaching documents in her office. The former live in "spread out piles" allowing her to see what's there. Teaching piles grow during the semester and get discarded at the end of the semester. At home piles of project related documents move between active documents on her kitchen table and shelves within "arm's

distance" during the day. Carsten maintains project related documents in plastic folders on his home office desk. Liz keep project piles on the coffee table in her living room and less active documents are placed in binders found on the floor in a corner. Steve, systematically organizes his project into piles on his office windowsills - the taller the pile the older the project. He will place a document at an off angle to distinguish the top active part from lower inactive documents. Journals live a particularly harsh life in Steve's office amassed in a pile with their mailing wrapping left intact. Only when he gets time to read will he un-wrap an issue. If an article fancies him he rips out the abstract page and adds it to the relevant project pile for later review and electronic download. The sad remains of the journal make their way to the recycling bin.

The other social scientists report comparable organizing schemes. For one, the piles closest to the desk are the most current and relevant. As focus of work changes so does the location of the piles. Most participants report disassembling or discarding piles after a project is completed. These types of document housekeeping are often happen after a significant delay.

Likewise, individual organizing schemes cut across our electronic file folders and email archives. No two members overlap in their classification and storage of electronic documents. The number of subdivisions varies greatly as well. Liz has more than 12 sub-folders in her project folder, Caroline in contrast two. Email folders seem to be less organized than file folders and come with fewer subfolders if any. Liz has no subfolders. Steve has two. In many instances the electronic folders work in much the same way as physical piles. The most recent documents crowd the top of the folder window organized by "date modified" as older files linger towards the bottom, out of view and thought.

Some participants use email or gmail folders as prime storage devices for documents. One senior faculty member always go to deleted mail or sent mail to retrieve old documents associated with distributed collaborations. Several other scholars reported slowly giving up on folders to organize documents and relying more and more on search functions to retrieve project documents on their desktop, email and shared repositories.

The mobile office bag: The work bag or backpack constitutes a particular central "folder" as it bridges the numerous locations, piles, and folders. During the mapping each of us proudly showed off our work bag and its contents. In those bags one finds beyond a laptop, documents requiring attention over the next two to four days depending on how often the owners return to their document dumping grounds. Caroline offers a particularly rich description of her bag:

"The content represents whatever I'm working on at the time within a week's timeframe. It could be teaching, research, administration, a dissertation. The bag has to have a certain weights and width. I try to keep it down to what is reasonably going to be dealt with within two to three day framework, because when I'm at home I don't always have to go into the office. So, it has to cover the days I might not have gone in. When we were working on the grant, I would have been carrying a printed draft, worked on it. Some of that might have been electronic."

Tracking the process of documenting work

The mapping data leaves a static picture of a rather heterogeneous and dynamic field of documents where individual participants apply comparable but fragmented storage and classification practices. If we want to understand working and organizing in technology dense environments we must track documenting practices over time. We set the mapping in motion. Tracking the process of documenting work revolves around four data gathering techniques:

1) Automated tracking of digital traces: We set up a system for the automated gathering of digital traces based on the analysis of how each scholar organizes documents related to their collaboration. Automated logs of email and shared repositories allow us to follow the flow of documents, who contributes to what documents, when and where. Individual electronic desktops offer more of a challenge due to widely different documenting practices and privacy concerns. Depending on the individual scholar's organizational scheme we install a program on their desktops to automatically track project related documents over time and save the results to an online repository.

One finds an rapidly expanding number of software tools offering a wide range of tracking capabilities from minute keylogging (e.g., http://www.refog.com/) and screen recordings such as Camtasia Studio (Ignatova and Brinkman 2007; Tang, Liu, Muller, and Drews 2006) to simply tracking changes in designated file folders (e.g., https://dropbox.com/). For the pilot study we applied Refog and Dropbox. The latter synchronizes the content of designated folders on your desktop with a similar folder installed on other devices. A web interface allows one to track changes to the folder over time. It also can give designated collaborators access to content in the particular folder on once desktop if they install a similar folder on their own desktop.

We then analyze these digital traces for process cycles and other major events. For instance, when do people tend to work on the project? How does content flow among collaborators? How quickly do they respond to other's inquires or postings? How do they engage different documenting practices for various types of work?

Following these procedures in the pilot study we found that documenting work tended to cluster around meetings, whether face-to-face or virtual, and externally or internally imposed project deadlines. Email traffic and changes to electronic folder content show a flurry of activity the day before a meeting, during the meeting and right after the meeting. These project meetings appear to serve as deadlines structuring people's temporal organizing of their work. People strive to accomplish promised tasks before the meeting and follow up on documents after the meeting. During meetings one or sometimes two members takes notes subsequently turning them into summaries and to-do lists. Others might look for articles and material discussed. More substantial writing activities such as research memos, conference papers or proposals colonize larger chunks spread across the work week. To date we cannot find any particular patterns in how we individually allocate those writing tasks but they seem to come in blocks of at least two to three hours.

One full professor organizes his distributed collaborations with two or three colleagues in a circular pattern. The participants take turn working on the paper or proposal until a set deadline. By that time the document responsibility moves to the next person in the group whether any changes were made to the document or not. This produces a slow and sequential organization of the work with a shared attention to deadlines and commitments.

The key-logging software Refog offered a closer analysis of individual documenting practices. Tracking our desktop use over several weeks suggests at least two types of documenting practices: *knowledge light* and *knowledge heavy* documenting work.

Knowledge light documenting work involves tasks that do not require our full attention, allowing us to easily move between different subtasks or let ourselves get distracted. Typical knowledge light activities include checking email, editorial, service and teaching activities.

Checking email constitutes a quintessential example. We tend to have multiple applications and documents open at the same time, constantly switching between them. For instance, an email may call our attention to a website or includes a document we need to review. That in turn requires us to create a new folder or revise yet another document. Meanwhile we often let ourselves get distracted by yet another mail popping up. Service and teaching work tend to depict the same patterns. When dealing with teaching issues Carsten typically has the school's web-based learning management system, multiple documents, email, PowerPoint and some pdf files going at the same time. Provenance plays a central role in this type of documenting work. Content is copied and pasted from one document to another or an older version of an assignment or presentation is edited, updated or otherwise modified to mach current needs. A recommendation letter may be written using an old one as a template prompting the faculty member to include certain elements or use particular verbiage. Student data are accessed in online repositories and desktop files and the recommendation letter is submitted in a web-based system.

Knowledge light activities do not appear restricted to particular places such as the office, but permeates most work locations. This is project coordination "busy" work, rarely requiring our undivided attention and we coast along drawing on previous instances of the same work.

Knowledge heavy work involving more reflection and thinking seems to call for other documenting practices. Often certain locations are disassociated with this type of work. Steve uses MS Word mainly when not in the office. Caroline and Carsten likewise rarely do any substantial writing in their work offices. Fewer applications and more physical documents are prevalent. Word and pdf dominate electronic applications and physical documents tend to include hand annotated articles and older drafts and scribbled note pads. Carsten and Caroline in particular spread physical documents across their desk and will produce outlines or rough drafts in hand. Caroline insists using a notebook for early phases of writing and thinking. She explains:

"If you can click text and stick it into your document that is easy, you don't have to go through it. But if you have to write it out by hand, first of all you take only what really matters, and I tend to do a better job writing up a summary or throwing in extra ideas. And I

can re-find my own ideas better on paper. Because all [electronic] text looks the same, there are no big strange arrows. . ."

2) Behavioral queries: We track the participants' activities through a number of behavioral queries, which repeat a variation of the virtual mapping method described above, including A) participant produced images, B) diary entries, and C) short interviews. Let us address them in turn.

Each participant is equipped with a digital camera and we ask them to take *pictures* of their desk tops, whiteboards or other tangible documents at select intervals throughout the study (Brown, Sellen, and O'Hara 2000; Murthy 2008). Repeating the mapping process detects shifts in how collaborators classify, store, and retrieve documents and identifies what documents get used or discarded over time. In general our experiences with images follow in the steps of other scholars. The main benefits of images come not from the photographs themselves, but from the participant's description of the actions associated with it and processes that lead to them taking that image (Brown, Sellen, and O'Hara 2000; Carter and Mankoff 2005).

We send out regular emails asking participants to respond the short questions. These have the character of *short diary entries* and track the unfolding coordination and document production. The diary method has received some attention recently by the information systems and human computer interaction fields. Several authors find it useful to study group information behaviors as it tends to capture rich data on participants' actions, interactions, feeling and reflections that otherwise would be difficult to access (Brandt, Weiss, and Klemmer 2007; Czerwinski, Horvits, and White 2004; Hyldegard 2006; Hyldegard 2008). However, participants quickly find regular diary entries to become a burden or simply forget. Some scholars have experimented with alternative methods such as having participants capture small "snippets" of information in the field via text, image, screen shorts, or voicemail. These bits then serve as prompts when participants eventually complete their diary at a more convenient time (Brandt, Weiss, and Klemmer 2007).

In our pilot study we are experimenting with a simple template in which participants can record in chronological order their documenting practices and briefly add reflections relevant to their task. The template provides a quick means for data entry and probes for details of participants, time and location, applications and documents used, the nature of the task/event, and a brief reflection on the process.

Short phone calls to individual scholars can supplement diary entries with *targeted interviews* around key project events. Alternatively, Palen and Salzman suggest using "voice-mail" diaries, where participants call in and record a message (Palen and Salzman 2002).

In the pilot study the ongoing behavioral inquires called out attention to how we in our collaboration relied little on centralized classification, storage and retrieval of documents. In other words, our coordination did not unfold around a central repository but relied on individual documenting practices with email being the point of contact and exchange. As with many of our previous collaborations we attempted a first to establish a shared repository

for project related documents. This process started with negotiations of what platform the group would use with individuals lobbying for their favorite technology and organizing scheme. Confusion or disregard for the shared space followed implementation. At first we established a Sharepoint site for the project (Microsoft product) and Liz our graduate assistant diligently saved project related documents to the site. The repository lives on largely untouched. Only on one occasion did a faculty member retrieve a document from the site. It also created some confusion as participants would not be sure whether or when to look for new submissions to the site.

Later we introduced Dropbox as a tracking tool for our documenting work over time. The site also gave us access to each other's electronic project related folders. This started to change our sharing practices. Instead of solely relying on email to share and coordinate documenting work we would at times access shared work through Dropbox. In the words of Liz:

"When I was tasked with analyzing each group member's work tracking chart I simply accessed them through Dropbox instead of having everyone individually email these charts as attachments. I did, however, have to remind one faculty member to put the document into Dropbox in the first place. I reminded him through email and face-to-face."

The participants agreed that those changes come slowly. Steve explains: *"I find myself more comfortable with the Dropbox being the pilot project folder. Administrative documents still reside on my local drive. I've not yet sorted out how to organize a shared space, so I drop everything into the pilot folder."* These emerging practices often leave participants in confusion. In fact, nobody interviewed for this research could explain how their collaborators organized their documents. Individual collaborators are left with the task of decoding their collaborators' classification schemes and storage practices: Where would Carsten have saved his memo? What interview transcript may be the most recent? After exploring his collaborators' files Carsten explains: *"Frankly, it seemed like a hodgepodge to me. It's hard for me to discern how they are used. I'm sure it's because it is not structured around my own specific practices."* As a consequence email remained the primary tool for sharing and coordinating documenting work. Participants circulate emails upon completing a documenting tasks and collaborators store and classify the document according to their own schema.

The other research subjects report similar stories about shared project repositories. Many have attempted or have established shared repositories using Learning Management Systems, Google mail and documents, or MS SharePoint. With the exception of one interviewee all report mixed experiences. Often a system is established but rarely, if ever, used. Typically, these repositories become a place for dead documents. In most cases there is no clear organizing scheme and nobody take charge organizing and keep up an organizing structure. A senior faculty member constitutes one exception to this pattern. She tends to be the driver for many of her distributed projects, which she conducts via email and a blackboard learning management site. She maintains a clear file structure for project documents and collaborators rely on her to push the project forward and keep documents organized.

3) Tracking key documents: Through the ongoing analysis of digital and behavioral traces we identify a small number of key documents playing a central role in the scholars' coordination activities. We track these documents in more detail and analyze how they evolve through different iterations. Content and document level network analysis helps us learn how key documents relate to other artifacts and events.

In our pilot study documents permeate all individual and collective project oriented activities among the four core and three peripheral members of the research group. Individually we read and write emails, edit proposal, find articles, write field notes and conference proposal, summaries of discussions and to-do-lists. When interacting synchronously documents directly frame or summarize the discussion. Asynchronous interactions takes place through emails, documents in their own right and often come with documents attached. In this way, the participants share a host of different documents including among others field notes, project proposal, pointers to articles and relevant web sources, virtual introductions to colleagues, conversation summaries, to-do lists, etc.

Among this plethora of documents two types stand out as the most prevalent: 'externally targeted documents' and 'to-do lists'. The externally targeted documents include grant proposals, conference papers, articles and other major textual products emerging from the collaboration. To date we worked on four grant proposal and five conference papers. Most of our coordination activities go into the production of these externally addressed documents. The to-do lists support the production of these major texts. They may merely be lists outlining activities to be accomplished by the group or specific members. For instance, when preparing for this pilot study Liz produced a simple table summarizing the methods we planned to apply. For several weeks this documents served as the pivot in our individual and collective work. In other instances, the to-do list takes a more elaborated form. Summaries of our weekly meetings constitute such lists as they capture the gist of conversations, steps taken and what activities lie ahead. When members share drafts of major written products they often attach a to-do list with activities left to be addressed. This can be unfinished sections, remaining literature searches or any other lingering work.

Both document types constitute *representations of* our work and *representation for* our work. The to-do lists squarely articulate this duality. One the one hand, the summaries and lists of activities represent the discussions and reflections that have gone into their production. One the other hand, they clearly offer a model for future activities in their outline of pending actions. The same dynamic plays out with the externally targeted documents. The early iterations of grant proposals and articles exchanged within the group most clearly show this dynamic. These first drafts often take the form of outlines with some sections holding snippets of text taken from previous work, others left empty. Again these outlines amount to a model of our thinking and a model for the work we have just started. In their later iterations and finished form externally target documents maintain the dual role. At the time of submission this article is a representation of our collaborative work to date. But in a few months' time it will be a representation for what I the author will have to do as I prepare for my EGOS presentation. As a group we are likely to use the paper as a model for what we

need to consider as we write our next grant proposal and publication as it add elements to the outlines of those future documents.

Not surprisingly the tracking of externally targeted documents and to-do lists becomes an important part of understanding work and organizing in our group's technology dense environment. Let us give a few examples: Our group's operational leadership appears to be closely tied to the person producing the to-do lists. For a long period, Liz, the graduate assistant, has been in charge of writing the to-do lists and circulating them by email. Her correspondence prompts the rest of us to engage in specific project activities, which she subsequently summarizes. Coordination leadership changes hands over time. For instance, when our group was approaching a major deadline for an externally targeted document, a grant proposal, the task of writing to-do lists shifted from the graduate assistant to Carsten, one of the faculty members as he started working on the first draft of the proposal. Periodically, Carsten would send out drafts accompanied by a to-do list. As the deadline drew close, Steve took charge of the proposal document and with it the role of writing to-do lists. After submitting the proposal the to-do lists authorship returned to Liz and with it, coordination leadership.

Among the other faculty members we are studying, several report similar circulations of coordination leadership. However, we also find variations. One senior faculty member reports that she is the driving force in all her virtual collaborations. Another senior faculty member structures coordination in a sequential order where each group member works on a paper in turn and transfers the responsibility to the next in line at a preset date. He keeps virtual projects small and offers the following explanation for this organizational scheme: "Limited admin needs."

As most other documents, the externally targeted documents and to-do lists do not reside in one place or remain in one materiality but move over time across folders, physical locations and materialities. If a document makes it into the mobile office bag it has attained a higher status in the project and receives a lot more attention. The most recent version of a grant proposal, for instance, initially played an important role framing and coordinating our pilot study, only to become obsolete after several weeks. It is no longer in the work bag that highlights people's center of attention. Likewise, a document's importance can be tracked as it gradually slides from a penthouse position at the top of a physical pile of documents to the bottom. Documents do not have equally long expiration dates. Externally targeted documents tend to remain central over extended periods. In contrast, collaborators engage frequently with the weekly summaries and to-do lists found on top of folders and piles but decline in use after the production of a new to-do list.

The location of emails also matters. The inbox seems to serve as people's individual to-do lists. We all leave "active" messages in the inbox as reminders to reply or finish tasks associated with the document. Messages with content that have longer term relevance are not only saved in the project's email folder but transformed into MS Word document, saved on the desktop, printed and added to the mobile work bag. The weekly to-do lists and many email attachments follow this trajectory.

For several interviewees tracking and organizing email threads take on particular significance. Steve, for instance, carefully edits email threads before saving a thread which encapsulates a full conversation. That way he can return and brush up on a discussion without having to go through several email messages in a subfolder. Participants in another large collaboration report that their email threads are "whale" shaped. A conversation tends to start with a few participants including in the address list, then moves on to include more if not all project participants, only to end with a few participants on the send/receive list closing the discussion.

Changes or new phases in coordination go hand in hand with the creation of new piles and folders. Carsten, for instance, created a file folder called "pilot study" when the collaboration moved from grant writing to testing the proposal ideas. As the content expands new sub-folders get added (e.g., analytical memos, interviews, etc.). This allows one to track project coordination as it evolves over time and its infrastructure expands. Steve explains: *"I'll create an ever-deepening set of sub-folders as more work gets done. My longest-living project has close to 50 sub-folders across four levels."* Two interviewees could not confirm such practices. They relied more and more on their computer's search function and would simply search their inbox, sent or delete email folders to find project relevant documents. Paper or project titles are sued as email subject headers to facilitate searches.

4) Focus group interviews: Scholars tend to meet face-to-face at conferences and workshops to discuss the progress of their distributed collaboration. At times they have phone-conferences. The groups we are studying tend to experiment with how best to conduct group meetings. Some combine co-located meeting with virtual participants. In those situations we tend to find that virtual members feel marginalized or shortcut by the technology. For instance, during one research group meeting a PhD student presented on his project progress. However, he had forgotten to circulate his slides prior to the meeting leaving the virtual members nothing to look at during the presentation broadcast via a phone conference system. Others keep all members virtual – even if some members share the same office building. Two interviewees reported that they rarely meet with local collaborators face-to-face but rely on email and a shared repository as if virtual colleagues.

When research groups meet we apply participant observation and focus group interviews to learn more about what meaning and understandings members associate with particular documents and the overall flow of their coordination in action. We have not conducted any focus group interviews among ourselves.

Discussion

Let us now return to our overarching methodological question: How do we best study documenting work in technology dense environments? Placing documents and the practices associated with them at the center of our qualitative methodology, as articulated above, comes with clear benefits, but also some new challenges. Table 1 summarizes the insights gleaned from our pilot study. We will discuss these in turn.

Benefits and Challenges	Comments
Holy trinity of qualitative methods	Multimodal gathering techniques are essential, where document mapping and tracking go hand in hand with interviewing and participant observation. The goals is to establish co-presence and not necessarily co-location.
Digital instrumentation	Software tools allow researchers to track a large range of document activities. To gather data with the appropriate granularity for addressing the research question requires the scholars to carefully select and combine their document tracking tools. Exploring socio-material arrangements becomes essential.
Documents as process/practice	The method allows researchers to study documents as processes or practices with temporal, spatial, and material manifestations that change over time.
Mapping versus flow	Mapping fields of documents is easier than tracking their flow. A hermeneutic approach that integrates repeated short cycles of data gathering and analysis brings document flow to the foreground.
Practicality	Many document tracking strategies require the research subjects' active engagement. Researchers must make sure that these do not become a burden on the subjects or simply neglected.
Privacy	Monitoring people's physical and electronic documents bring the researcher in proximity with highly private information. Depending on the subjects documenting practices data gathering instruments must be selected and adjusted to safeguard their privacy. Enabling the subjects to control the monitoring and review the data before released to the researcher further protect subjects' privacy.
Heisenberg-style challenge	A document centric methodology turns the research subjects' attention to the infrastructure of their documenting work and they often start altering their practices if only temporarily. Such affects need not reduce the reliability of the method but can bring practical and conceptual insights to a study.

Table 1: Methodological benefits and challenges

Holy trinity of qualitative methods: As suggested by our pilot study documents do not have to live as second class citizens in our qualitative methodologies. They offer a window into collective practices and organizing in technology dense environments. One can approach documents not only as artifacts but also utterances and traces of practices (individual and collective) in line with data produced through interviews and participant observation. Hereby, we do not suggest that the researcher stop being a central data gathering 'instrument' and revert to more mechanical and de-contextual data gathering techniques, such as, survey studies or double blind experimental setups. Interviews and participant observation remain central methodological tools in our pilot study. In fact we regard documents, interviews and participant observation as the holy trinity of qualitative methods. One does not go without the others.

If we accept this trinity – the important question becomes how we combine the three (i.e., document collection, interview and observation) to create *co-presence* as opposed to merely co-location (Beaulieu 2010). For most ethnographers and qualitative researchers the paramount method question has been: where do I go? The space shared with research subjects has been the main source of data for many qualitative researchers. A reliance on co-location is not viable in technology dense environments. One can simply not restrict the social organization of work to a limited set of locations which a researcher can practically inhabit. A

document centric research instead strives for co-presence which can be established through multiple modes of interaction. Some might involve email exchanges, listening in on phone conferences, or looking through digital traces left in repositories such as Dropbox, discussion boards, or analyzing the changes in folder structures in a shared repository. Mapping and tracking documenting work becomes an important way of establishing such co-presence in combination with interviews and participant observation in person or virtually.

The possible combinations of interviews, documents and observation should of course be considered in light of the overall project and question to be explored (Taylor 1999a). For technology dense environments we find that exploring the technical and social boundaries people live in require multiple modes of data gathering where one engage with participants on-line, off-line and through other modes of interaction characteristic for the setting.

 We saw in the pilot study that document gathering and analysis can prepare the researcher for interviews and participant observation. As ethnographers may engage in weeks or months of observation before doing their first interview, documents often constitute a necessary foundation on which to conduct informed interviews. We find that one or more documents collected from the field may become an interview guide. Instead of preparing a set of semi-structured interview questions documents can give structure to an interview and become the artifact around which the conversation evolves. The reverse relationship is also evident. Interviews and participant observation are necessary tools to set up a reliable and valid way to gather documents over time. One cannot build a system for collecting and tracking documenting practices without the initial mapping of each participant's field of documents, where interviews and participant observation play an important role.

Digital instrumentation: Mapping and tracking people's documenting practices adds a layer of practical concerns to a qualitative study, which in some ways compare to the way bench sciences must deal with instrumentation. As noted, numerous programs exist to collect digital traces including software that simply tracks changes to specific folders (e.g. Mozy and Carbonite Pro) to tools that capture all the user's desktop activities such as screen recording software like Camtasia Studio (Ignatova & Brinkman 2007; Tang et al 2006) or keylogger software like Refog and Spytector. Some communities and professions have tailored tools to track document changes, open source programmers being an obvious example (e.g., http://mercurial.selenic.com/). Given such a broad range of ever-expanding options, determining the appropriate software needed to capture data with the desired level of granularity becomes an important task.

How best to do this depends again on the project and its research questions. It is an integral part of one's research design. However, it does not have to be cast in stone as if it were a laboratory setup. As ethnographers has been adjusting their participation and interview techniques to the context and the evolving relations to informants, so does a document centric study call for a flexible approach to data gathering where the researcher continuously adjusts and refines the data gathering approaches based on what has been learned to date.

In technology dense environments, such adjustments require continuous exploration of materiality. Just as any participant might be hard to nail down or track, documents are easily lost, overlooked, or overwritten. Using various techniques to track down these artifacts is often required (Taylor 1999b). Reliance on file libraries, document archives, bulletin board and the like become very useful in building complete account of a technology dense environment. Keeping up with new tools which allow the participants to work and organize and the research to map and track documenting work are likely to pay off in the long run.

In our pilot study Dropbox (http://www.dropbox.com) has proven particularly helpful in monitoring the history of document revisions by creating a log of when documents were created, modified, and deleted by participants. Software add-ons offer additional functionality, such as screen capture and mail drop, a feature which automatically downloads and saves email attachments to the desktop. While generally unobtrusive, Dropbox requires minor alterations of participants' work practices. For example, a document file can only be monitored if it is directly placed into the Dropbox folder located on the participant's desktop. In contrast, keylogger software such as Refog captures a plethora of information such as keystrokes, web browsing history, screenshots of applications and files without changing the participants' organizing behavior. The raw data accumulated from Refog while vast is also searchable and makes the volume of data collected more manageable. Inevitably, key-logging software raises some serious ethical concerns and will likely produce more data than one would want.

Email tracking offers another challenge. Distributed collaborators often use different email clients and operating systems which limits the types of email tracking software that can be installed on participant computers. Key-logging software such as Refog proves inadequate to capture email and in particular text heavy communication. It is possible to configure each participant's email client to automatically copy a research account on all sent messages but it requires the researcher to routinely filter out email not pertaining to the research study. Alternatively, one can task the research subject to manually copy a research account on all project relevant emails.

Documents as process/practice: Conceptualizing documents, not as stable artifacts but processes or practices with temporal, spatial, and material manifestations that change over time are central to our document-centric methodology. The documents studied in the pilot emanate from, if not constitute the work practices associated with our small social science collaboration. We noticed that some interviewees engage with colleagues in the same build as they do with colleagues at neighboring continent. Without studying the documents it would be difficult to comprehend the collective activities and organizing.

One of the benefits of the methodology outlined above is that it directs attention to the temporal, spatial, and material dimensions of documenting practices. The documents studies in the pilot changed over time. The content may change but equally important their spatial location and materiality shift with significance for the individual and collective actions studied. For instance, tracking the specific location of documents in a physical pile or whether a document could be found in the office or work bag matters and offers insights to

the unfolding coordination of our group. Do we bother to file a document we receive as an email attachment on our desktop or do we shove it in an email folder to be forgotten? Changes in materiality hint at similar changes in a document's importance for our group's current activities. We do not bother to print out all documents and lug them around in our office bag or spread them around our keyboard on our desktop. Each of these changes captured by a documents-centric methodology serves as snapshots of the unfolding work and organizing in technology dense environments.

A comparable strategy can be found in science studies and in particular among actor network inspired scholars. Jensen and Lauritsen (2005) suggest that we 'read with documents' as opposed to 'reading against documents.' The latter approach develops out of discourse analysis and focuses on a critical reading to uncover hidden, forgotten or repressed meanings and make such themes explicit. Reading with the documents, in contrast, involves traveling with the document to see where it goes and what relations get established or severed on the way. Latour and Woolgar's (1986) classic lab study likewise track documents through the laboratory and beyond as a way to understand the production of knowledge in those technology dense laboratories.

Mapping versus Flow: During the pilot study we found that mapping a field of documents was easier than tracking the flow of documenting work. The tracking process required new research practices to fully bring our attention to the flow and sharing of documents, while mapping lends itself easily to more traditional interview and participant observation techniques which are familiar to most social scientist. One can go out to interview and observe one participant at a time, return home and develop a nice description of the informants' field of documents and the meanings they associated with various documents and the practices that go into their production, classification, storage and retrieval.

To study flow requires the collection and intermediate analysis of larger bodies of documents before a picture emerges. Only then do interviews and observations become helpful in elaborating the dynamics of how documents are shared and link people's individual fields of documents. We found that a hermeneutic approach supports the tracking of documents by organizing the research into short repeated cycles of data gathering and analysis. The repeated mapping illustrates this strategy where the researcher performs an initial mapping of the document field followed by analysis which helps set up automated tracking and repeated behavioral queries (i.e., mappings). Analyzing the collected and tracked documents in between each behavioral query becomes necessary if one wants to illuminate the ongoing flow. Not only does it allow the researchers to refine their tracking techniques and time their mapping to important events and places for collaboration they study, it also makes it possible to triangulate document based observations through interviews, participant observation and new document analysis.

Practicality: Many documenting tracking techniques require the active involvement of the research subjects. Manually copying a research account on project relevant emails, writing diary entries on a regular basis, saving project documents in folders accessible to the researchers all consume the subjects' time and attention. Many computer science studies of

document provenance call for subjects to log changes to documents and applications over time (Dragunov, Dietterich, Johnsrude, McLaughlin, Li, and Herlocker 2005; Lonsdale, Jensen, Wynn, and Dedual 2010). From the research subjects' perspective these quickly become a burden or the tracking activities simply fall to the wayside as more pressing concerns crowd out people's attention and they simply forget to write the next diary or cc the project account on an important email. Even when studying ourselves we found that keeping up with regular diary writing slowly fizzled out over the course of a few weeks. Putting together a set of tools that weigh comprehension against the practicality of the data gathering from the research subjects' perspective will require critical decisions in most document centric studies. After all it is much easier to dream up elaborate data gathering techniques than implement them.

Privacy: Qualitative researchers have over the years developed a series of techniques to afford confidentiality to their research participants, and to deal with how to keep private participants' identity and behaviors. However, dwelling in somebody's document universe raises a number of additional privacy issues. Most people use their computers for multiple purposes, many reaching deep into their private sphere. It can be difficult to set up a system that monitors specific documenting practices only. As mentioned earlier, Dropbox and other file sharing systems allow one to track designated folders, but leave many activities invisible such as applications and documents used during the production, classification, storage and retrieval of relevant documents. Key-logging and screen capturing tools provide a wealth of data, yet their invasive nature makes them less palatable for IRB officers and participants alike (Tang, Liu, Muller, and Drews 2006). In our pilot study we addressed this inherent privacy concern by: 1) giving participants the control to pause and restart key-logging software and select among several tracking functions (e.g., screen shots and application tracking but not key-logging) at any time during the study; 2) allowing participants the option of deleting any data listed in the software's activity log. We also experimented with giving the subjects the option of analyzing their own tracking data and reporting patterns back to the research.

With an increased attention paid to information privacy legislation even less intrusive methods call for careful attention. A number of backup and file sharing software tools allow researchers to track only specific folders or items on the research subjects' computers and save it in online repositories easily accessible to the researcher. Choosing password protected and encrypted services will in most situations be advisable.

Heisenberg-style challenge: All qualitative research methods face a Heisenberg-style challenge where the process of observing shapes the observed. Certainly this was the case when we turned our own practices into the focus of the pilot-study. We found that when one starts asking people about their documenting work they turn their attention to the infrastructure of their work and inevitable start to clean it up. This can include straightening piles of documents around the office, deleting icons on a cluttered computer desktop, or finally taking the time to clean out a bulging email inbox. Some of us reinforced classification and storage schemes in which we believed but rarely practices. The urge to tidy up one's documents may at times be attributed to a sense of "airing one's dirty laundry" in

front of strangers. As socio-technical artifacts, documents reveal a great deal about not only a technically dense environment, but also idiosyncrasies and personal quirks.

Instead of seeing this as a threat to the reliability of the method it offers important insights into how research subjects perceive and work with their document infrastructure. How malleable are their documenting practices and how close a fit one finds among articulated documenting principles versus the reality of unfolding documenting practices. Through repeated behavioral queries, and triangulation among participants, researchers can deepen their understanding of whether such changes survive over time or return to a previous equilibrium. For instance, the email inbox may only stay organized for a few days past the first observation. Returning to the subject repeatedly will allow the researcher to verify such observations. Thus, having research subjects react to the process of observation and the insights it gives them into their own documenting practices need not reduce the reliability of a document centric methodology, but may enhance it; rather than affecting the subjects too much, deliberately encouraging the research subject to react to the research data and observations may strengthen the methodology.

In short, issues of verifiability and reliability touch all research projects – document centric studies being no difference. Recognizing and using these issues to a project's benefit is an important part of any study. Triangulating findings using multiple data sources always helps. Exploring how particular sociomaterial assemblages shape relations among subjects and researchers becomes equally important to further develop a document infused research method.

Conclusion

To illustrate our document-centric methodology to the study of distributed work in technologically dense environments we draw from our ongoing study of virtual organizing among distributed social scientists. Over the past two years our own and ten colleagues' virtual organizing as served as a methodological test bed where we could try out various strategies and techniques. These data allows us to demonstrate the benefits and challenges involved in studying work practices in technologically dense environments through people's documentary practices.

The document-centered methodology developed in this paper offers a new perspective on work and organizing in technology dense environments and a call for revised research strategies. In and of themselves, participant observation, interviews, content analysis and other quantitative techniques do not offer a nuanced picture of technology dense environments. We find a need for first, an initial mapping of documents in the collaborative field before we start to track documents and behaviors over time. Second, the ongoing flow of virtual organizing only becomes apparent by triangulating the digital flow of documents, observation of tangible documents (e.g., paper) and repeated behavioral inquires. Here, it is important to be sensitive to the rhythm of the collaborative work studied. Inquires about documenting work tend to produce the most interesting responses if questions are timed closely with their production and use. Third, the ongoing research suggests that documents

supporting work practices do not serve as stable information artifacts. Rather they become snapshots in time, part of the general flow of coordination in action across numerous documents and applications.

The research strategy offers some insights into social scientists distributed organizing including our research group's work during the pilot study. We did not find any stable infrastructures or formal preplanned arrangements that embody the design of these small organizations. Rather the organizing unfolded around the production of a number of documents, some of which mainly supported our unfolding practices, such as to-do-lists and summaries, while others served as externally targeted documents. We did not have any overarching classification, storage and retrieval system for these shared documents. Attempts to implement such centrally organized sharing schemes largely led to non-use or confusion. Instead the documents themselves served as the linking pins among our distributed practices. Most participants would maintain their own classification, storage and retrieval systems which supported their particular work habits. Shared systems were only viable as long as they supported these individual practices. Email largely served as the vehicle for document sharing.

The circulated documents constituted the main point for organizing as they offered representations *of* our current thinking and activities and representations *for* what to do next. Sometimes those prescriptions were explicitly articulated in to-do lists specifying who was responsible for what activities and when. In many other cases, the basic structure of the documents (e.g., its outline) or genre expectations helped structure individual practices.

These documenting practices didn't constitute a stable infrastructure. Rather, they changed over time as the work evolved, deadlines approached, or tasks shifted. For instance, the coordination leadership which was tied to the production of to-do lists changed hand over the course of the project circulating among members depending on the type of activities currently the focus of the group. Individual documents likewise did not serve as stable coordination artifacts. An externally targeted document may start out as a mere outline, specifying the work to come. As the document was finalized it became a product of the group, marking the end of a coordination phase, only to become the main point of reference a few weeks later as the collaborators started to work on a new article.

The particular material manifestations of those documents, their location in space and time offered important insights to our organizing and work. If a document circulated among us was transformed from an email attachment to a paper document and placed in our work bag it indicated that this document was going to organize our practices in the near future. Likewise, if a new folder or pile was instigated by a participant it suggested a new phase in the collaboration. These changes did not bring collaborators into lock step coordination. Instead it is likely that only one or two participants may print out the document and carry it with them or create a new folder. For other participants the activities tied to the document may simply not be urgent or their involvement not required to a degree that warrant changes to the materiality of the document, it location and file folder structure.

Limitations: These insights should clearly be taken for what they are, a pilot study of our own and a few colleagues work and organizing in technology dense environments. The research was mainly carried out to develop and test a document-centric methodology. Future research will be required to further develop a significant understand of social scientists organizing in distributed and technology dense environments. We are in the process of organizing and establishing funding for a larger study which compares documenting practices across multiple virtual social science collaborations. Bringing such a perspective to other types of organizations in other industries would likely add other perspectives to how various professional groups apply documents to coordination their work.

In summary, the present paper speaks to one of the core questions associated with work and organizing in technology dense environments. How do technological and organizational practices interweave? The paper focuses on the role of documents in supporting distributed work among social scientists. Documents include both traditional paper-based material forms such as memos, reports, and publications, as well as online documents, web pages, and text-based communications such as email. For the casual observer, the omnipresence of documents and their myriad forms are taken for granted – a seemingly innocuous sidelight to doing science. Such a naïve view belies the importance of documents. For students of technology dense environments, documents can reveal the ways in which knowledge work gets done. The intent is to build on the more than 100 years of research in this area to advance conceptualizations of and methodological approach to current documentary practice in the age of distributed coordination.

References

Ackerman, M. S. and Christine Halverson. 2004. "Organizational Memory: Processes, Boundary Objects, and Trajectories." *Computer Supported Cooperative Work* 13:155-189.

Bahtia, V.K. 1993. *Analysing Genres: Language use in professional settings*. London: Longman Group.

Bakhtin, Mikhail Mikhailovich. 1986. "The Problem of Speech Genres." Pp. 60-102 in *Speech Genres and Other Late Essays: M.M. Bakhtin*, edited by C. Emerson and M. Holquist. Austin: University of Texas Press.

Bakhtin, Mikhail Mikhailovich. 1996. *The Dialogic Imagination: Four essays by M.M. Bakhtin*, Edited by M. Holquist. Translated by C. E. a. M. Holquist. Austin: University of Texas Press.

Barad, K. . 2003. "Posthumanist performativity: Toward an understanding of how matter comes to matter." *Signs: Journal of Women in Culture and Society* 28:801-831.

Baudoin, P. 2008. "The principle of digital preservation." *The Serials Librarian* 55:556-559.

Bazerman, Charles. 1995. "System of genres and the enactment of social intentions." Pp. 79-104 in *Genre and the New Rhetoric*, edited by A. Freedman and P. Medway. London: Taylor & Francis.

Beaulieu, A. 2010. "From co-location to co-presence: Shifts in the use of ethnography for the study of knowledge." *Social Studies of Science*:1-18.

Blanc-Brude, T. and D. L. Scapin. 2007. "What do people recall about their documents? Implications for desktop search tools." Pp. 102-111 in *the 12th International Conference on Intelligent User Interfaces*. New York: ACM.

Brandt, J., N. Weiss, and S. Klemmer. 2007. "Txt 4 l8r: Lowering the Burden for Diary Studies Under Mobile Conditions." in *ACM, CHI 2007*. San Jose, CA, USA.

Briet, Suzanne. 1951. *Qu'est-que la Documentation?* Paris, France: Editions documentaires, industrielles et techniques.

Briet, Suzanne, Ronald E. Day, Laurent Martinet, and Hermina G. B. Anghelescu. 2006. *What is documentation? : English translation of the classic French text*. Lanham, Md.: Scarecrow Press.

Brown, B., A. Sellen, and P. O'Hara. 2000. "A Diary Study of Information Capture in Working Life." in *ACM, CHI 2000*. The Hague, Amsterdam.

Brown, J. S. and Paul Duguid. 1994. "Borderline Issues." *Human-Computer Interaction* 9:3-36.

Buckland, Michael K. 2007. "Northern Light: Fresh Insights into Enduring Concerns." Pp. 327-334 in *A Document (Re)turn: Contributions from a research field in transition*, edited by R. Skare, N. W. Lund, and A. Varheim. Frankfurt am Main: Peter Lang GmbH.

Carlile, Paul Reuben. 1997. "Understanding Knowledge Transformation In Product Development: Making Knowledge Manifest Through Boundary Objects." Unpublished Ph.D. Dissertation Thesis, University Of Michigan, USA.

Carlile, Paul Reuben. 2004. "Transferring, Translating and Transforming: An integrative framework for managing knowledge across boundaries." *Organization Science* 15:555-568.

Carter, S. and J. Mankoff. 2005. "When Participants Do the Capturing: The Role of Media in Diary Studies." in *ACM, CHI 2005*. Portland, Oregon.

Crowston, K. and B. H. Kwasnik. 2003. "Can document-genre metadata improve information access to large digital collections?" *Library Trends* 52:345-361.

Czerwinski, M., E. Horvits, and S. White. 2004. "A Diary Study of Task Switching and Interruptions." in *ACM, CHI 2004*. Vienna, Austria.

Dragunov, A.N., T.G. Dietterich, K. Johnsrude, M. McLaughlin, L. Li, and J. Herlocker. 2005. "TaskTracer: A Desktop Environment to Support Multi-tasking Knowledge Workers." Pp. 75-82 in *IUI'05*: ACM Press.

Forsythe, Diane. 1999. "It's just a matter of common sense: Ethnography as Invisible Work." *Computer Supported Cooperative Work (CSCW): An International Journal* 8:127-145.

Frohmann, B. 2007. "Multiplicity, Materiality, and Autonomous Agency of Documentation." Pp. 27-39 in *A Document (Re)turn: contributions from a research field in transition*, edited by R. Skare, N. W. Lund, and A. Varheim. Frankfurt am Main: Peter Lang GmbH.

Frohmann, Bernd. 2004. *Deflating information: From Science studies to documentation*. Toronto: University of Toronto Press.

Geiger, S. and D. Ribes. 2011. "Trace Ethnography: Following coordination through documentary practices." in *The 44nd Annual Hawaii International Conference on System Science (HICSS-44)*. Hawaii, HI: IEEE Computer Society Press.

Harper, Richard. 1998. *Inside the IMF: An ethnography of documents, technology, and organizational action*. San Diego: Academic Press.

Hine, C. 2006. *New Infrastructures for Science Knowledge Production: Understanding E-Science*. Hershey, PA: Idea Group.

Hyldegard, J. 2006. "Using Diaries in Group Based Information Behavior Research-A Methodological Study." in *ACM Information Interaction in Context IIiX*. Copenhagen Denmark.

Hyldegard, J. 2008. "Beyond Search Process-Exploring Group Members' Information Behavior in Context." *Information Processing and Management* 45:142-158.

Ignatova, E. and W. Brinkman. 2007. "Clever Tracking User Behavior over the Web: Enabling Researchers to Respect the User. ." in *21st British Computer Society HCI Group Conference*, vol. 2. Lancaster University, UK.

Jensen, Casper Bruun and P. Lauritsen. 2005. "Reading Digital Denmark: IT-reports as material semiotic actors." *Science, Technology and Human Values* 30:352-73.

Knorr Cetina, Karin. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge: Harvard University Press.

Latour, Bruno. 1987. *Science in Action: How to follow scientists and engineers through society*. Cambridge, Mass.: Harvard University Press.

Latour, Bruno and S. Woolgar. 1986. *Laboratory Life: The construction of scientific Facts*. Princeton: Princenton University Press.

Law, J. 2004. *After Methods: Mess in Social Science Research*. New York: Routledge.

Lonsdale, H., C. Jensen, E. Wynn, and N.J. Dedual. 2010. "Cutting and Pasting Up: "Documents" and Provenance in a Complex Work Environment." in *The 43nd Annual Hawaii International Conference on System Science (HICSS-43)*. Hawaii, HI: IEEE Computer Society Press.

Lund, N. W. 2009. "Document Theory." *Annual Review of Information Science and Technology* 43:399-432.

Lynch, C. A. 2001. "When documents deceive: Trust and provenance as new factors for information retrieval in a tangled web." *Journal of the American Society for Information Science and Technology* 52:12-17.

Miller, Carolyn. 1984. "Genre as social action." *Quarterly Journal of Speech* 70:151-167.

Murthy, D. 2008. "Digital Ethnography: An Examination of the Use of New Technologies for Social Research." *Sociology* 42:837.

Orlikowski, W. J. 2007. "Sociomaterial Practices: Exploring technology at work." *Organization Studies* 28:1435-1448.

Orlikowski, Wanda J. and JoAnne Yates. 1994. "Genre repertoire: The structuring of communicative practices in organizations." *Administrative Science Quarterly* 39:541-574.

Østerlund, C. 2007. "Genre Combinations: A Window into Dynamic Communication Practices." *Journal of Management of Information Systems* 23:81-108.

Østerlund, C. 2008. "Documents in Place: Demarcating Places for Collaboration in Healthcare Settings." *Computer Supported Cooperative Work (CSCW)* 17:195-225.

Østerlund, C. and D. Boland. 2009. "Document Cycles: Knowledge Flows in Heterogeneous Healthcare Information System Environments." in *The 42nd Annual Hawaii International Conference on System Science (HICSS-42)*, edited by J. F. Nunamaker, Jr. and R. H. Sprague, Jr. Hawaii, HI: IEEE Computer Society Press.

Otlet, P. 1903. "Les sciences bibliographiques et la documentation [Bibliographic science and documentation]." *IIB Bulletin* 8:125-147.

Otlet, P. 1907. *L'organisation systematique de la documentation et de la developpement de l'Institut International de Bibliographie [The systematic development of documentation and development of the International Institute or Bibliography]*, vol. IIB Publication No. 82. Brussels, Belgium: The Institut.

Palen, L. and M. Salzman. 2002. "Voice-Mail Diary Studies for Naturalistic Data Capture Under Mobile Conditions." in *ACM, CSCW '02*. New Orleans, Louisiana, USA.

Pollock, N., R. Williams, and . 2010. "e-Infrastructures: How do we know and understand them? Strategic ethnography and the biography of artefacts." *Computer Supported Cooperative Work* 19:521-556.

Randall, Dave, Richard Harper, and Mark Rouncefield. 2007. *Fieldwork for design: Theory and practice*. London: Springer.

Shen, J., E. Fitzhenry, and T. G. Dietterich. 2009. "Discovering frequent work procedures from resource connections." Pp. 277-286 in *the 13th International Conference on Intelligent User Interfaces*. New York: ACM.

Star, S. L. and J. R. Griesemer. 1989. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology 1907-39." *Social Studies of Science* 19:387-420.

Star, S.. L. and K. Ruhdeler. 1996. "Steps toward an ecology infrastructure: Design and access for large information spaces." *Information Systems Research* 7:111 - 134.

Suchman, L. A. 1987. *Plans and Situated Actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Suchman, L. A. 2000. "Making a Case: Knowledge and routine work in document production." in *Workplace Studies: Recovering work practice and informing system design*, edited by P. Luff, J. Hindmarsh, and C. Heath. Cambridge: Cambridge University Press.

Suchman, L. A. . 2007. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press.

Swales, J. 1990. *Genre Analysis*. Cambridge: Cambridge University Press.

Sweeney, S. 2008. "The ambiguous origins of the archival principle of "provenance." " *Libraries & the Cultural Record* 43:193-213.

Tang, J., S. Liu, M. Muller, and C Drews. 2006. "Unobtrusive But Invasive: Using Screen Recording to Collect Field Data on Computer-Mediated InteractionScreen Recording to Collect Field Data on Computer-Mediated Interaction." in *CSCW*. Alberta, Canada.

Taylor, T.L. 1999a. "Life in Virtual Worlds: Plural existence, multi-modalities, and other online research challanges." *American Behavioral Scientist* 43:435-449.

Taylor, T.L. 1999b. "Life in Virtual Worlds: Plural Existence, Multimodalities, and Other Online Research Challenges." *American Behavioral Scientist* 43:436-449.

Yates, JoAnne and Wanda Orlikowski. 2007. "The PowerPoint Presentation and Its Corollaries: How Genres Shape Communicative Action in Organizations." in *Communicative Practices in Workplaces and the Professions: Cultural Perspectives on the Regulation of Discourse and Organizations*, edited by M. Zachry and C. Thralls. Amityville, NY: Baywood Publishing.

Zacklad, Manuel. 2006. "Documentarisation Processes in Documents for Action (DofA): The Status of Annotations and Associated Cooperation Technologies." *Computer Supported Cooperative Work* 15:205-228.