

# Actually Digital: Personal Electronic Libraries and a Reconstruction of Academic Reading Practices

Hilah Geer



European Science, Society, & Technology  
Master's Thesis  
The Center for Technology Information and Culture

UNIVERSITETET I OSLO

March 2012

*Word count: 20240*



# **Actually Digital: Personal Electronic Libraries and a Reconstruction of Academic Reading Practices**

© Hilah Geer

2012

Actually Digital: Personal Electronic Libraries and a Reconstruction of Academic Reading Practices.

Hilah Geer

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

IV

# Abstract

This thesis studies academic reading practices through the use of, Docear, a new software program with a module that combines PDF and mind mapping technologies. Five in-depth Skype interviews designed to reveal situated uses of Docear were conducted. The author's experience as an academic librarian serves in investigating the surrounding "infrastructure of relations" and has a reflexive role in the analysis. Actor Network Theory (ANT) and Science and Technology Studies (STS) form a theoretical and methodological basis. The analysis looks at digital libraries from a perspective of things virtual (potential) versus actual. Findings reveal new technologically-assisted practices related to keeping, annotating, reading, and structuring digital literature that employ visual mapping and where space is replaced by connection. The fragmentation of text and need for structure in academic reading practices are seen from within of a changed physical arena. Credence is given to the potency of making academic publications compatible to a mind mapping program like Docear.

Keywords: Science and Technology Studies, Actor Network Theory, Academic Reading Practices, Docear, PDF technology, Freeplane, Mind mapping, Information Science, Virtual vs. Actual.



In loving memory of

A. Katie Geer  
and  
William D. Geer, Sr.



# Acknowledgements

First I would like to thank Jöran Beel for being so open and responsive and making this happen without getting in my way in any way. I wish him and his colleagues best of luck with this fantastic project. And to the researchers who agreed to be interviewed, all of whom were so generous with their time. They had such useful and articulate ways of describing their work that at times it seemed it would be best to get out of their way and just publish their transcribed words. If they read this I hope they feel that what they shared of themselves is treated with care and that my own words did not get in the way of the descriptions of their experiences.

To my advisor Beate Elvebakk, thanks for a light touch, positivity, and steering me towards, among other things, Christine Hine, Lucy Suchman, and the idea of practice. And perhaps, most significantly, for telling me not to think too much until I had the interview text in hand. I also learned from and was inspired by the beautiful clear writing style of your PhD Thesis. Of course, that does not make you responsible for where I do not achieve the same here.

I would like to thank Kristin Asdal for a brief but encouraging chat about the virtual. Thanks to the TIK faculty and program in general. You have given me a new way of seeing things. Thanks to Ole Ronny Tveite-Strand and Øystein Moen for making the degree process work. Thanks to all my fellow students for helping me learn. I wish I had had more time to socialize.

Thanks to The Oslo School of Management and Campus Kristiania, specifically Line Christoffersen, for supporting my taking a year's leave to study. Thanks to my colleagues in the library: Live Vikøren, Hallvard Mildestveit and Urd Hertzberg. And thanks goes especially to Urd for taking over while I was away.

Thanks to Trude, Karl Georg and family who took care of me in the final throws-- unbelievable warmth, generosity, friendship, and moonshine on snow.

Thanks to my Mom and Dad and sisters Gretchen and Lyedie. You are all on every page in one way or a hundred others, some just more precisely.

And the giant thanks goes to Steinar, Katie Lou, and William Coley for just about everything and putting up with just about everything. Miss you!! Be back soon!!!



# Table of Contents

1	Introduction .....	1
2	Intentions and Context .....	5
2.1	Choice of Subject.....	7
2.1	Notes on My Approach to Writing Ethnography .....	11
2.2	Choosing an Object .....	12
2.3	Description of Docear.....	15
3	Micro Environments, Infrastructure, and Situated Use (Theory).....	21
3.1	The Micro & Infrastructure in Actor Network Theory.....	21
3.2	Ethnomethodology: Prototypes and Situated Use .....	24
4	Ontology of Actual & Virtual .....	29
5	Methods.....	33
5.1	Interview Process.....	35
5.2	Summary of the Interviewee Relevance .....	38
6	Findings and Analysis .....	43
6.1	Invisible Reading and the Lone User .....	43
6.2	Distilling Actual Digital Libraries .....	47
6.3	Making Actual Digital Libraries “Your Own” .....	52
6.4	Visually Structuring Actual Digital Libraries .....	54
6.5	Space as Connection in Actual Digital Libraries.....	59
6.6	Tangibility of Actual Digital Libraries.....	63
7	Conclusion.....	67
	List of References.....	71
	Appendix 1: Template for Introductory Letter and Terms of Agreement.....	74
	Appendix 2: Interview Script .....	75
	Appendix 3: Scribie’s Transcription Process .....	77
	Appendix 4: Call for Interviewees on Docear Blog.....	78

# List of Figures

Figure 1: Docear’s Presentation of their Academic Literature Suite .....	17
Figure 2. Expanding and Contracting Mind maps. ....	19
Figure 3: Summary of Data about Interviewees (Context and Relevance to the Study) .....	34
Figure 4: Summary of Findings: Actualizing Digital Literature in Docear .....	42
Figure 5: Distillation of Actual Digital Text.....	50
Figure 6: Working with an Actual Digital Library in Docear.....	57





# 1 Introduction

How do we pack the world into words? (Bruno Latour 1999, p. 24)

~

Where clarification is needed, the readers of a text cannot readily ask the authors what they meant. The focus in consuming texts is therefore placed far more on the interpretive work done by readers and less on the shared understanding between authors and readers. We tend now not to see texts as transparent carriers of the meanings intended by their authors. (Hine, 2000, p. 50)

Machine as text. (Woolgar, 1991, p. 61)

~

In so far as it is successful, the prototype works as an exemplary artifact that is at once intelligibly familiar to the actors involved, and recognizably new. ...

Prototyping practice simultaneously recovers and invents work requirements and technological possibilities, that each make sense in relation to the other. (Suchman, Trigg and Blomberg, 2002, p. 1)

This thesis studies changes in academic reading practices. Through interviews with five researchers, it employs a micro approach to build up a new perspective on work related to the consumption of academic text. The above pairs of quotations frame the subject and situate it in the field of Science & Technology Studies (STS), and more specifically its sub-branch Actor Network Theory (ANT). The first two quotes juxtapose the challenge of revealing how one writes scientific literature, well-studied by STS, with the lesser-studied and difficult-to-observe topic of the reading of scientific literature. The latter being how we unpack the world from words? The second pair of quotes says that because machines carry meaning they can, like a text, be read, but that neither machine nor meaning will come from

one static vision. A machine's design and its usefulness will always have a basis in something already recognizable. And if it affects change it will in turn be changed. This research occurs within this dynamic as it relates to changes in reading practices by users of a digital tool for personal library management.

The tool, Docear<sup>1</sup>, was used intensively by all 5 interviewees for their research. It calls itself an "academic literature suite" and has three modules. The module studied here assists in the personal management of electronic literature and tasks related to storing, retrieving, reading, organizing, annotating, and structuring academic text. This is an ethnographic study based on the author's in-depth Skype interviews with the users. Descriptions of situated use and perceptions of change were collected and analyzed. Use revolves around Docear's mind mapping features for creating and working with personal libraries of PDF files. Based on Actor Network Theory, observations are made and contingencies revealed within the dynamic network of humans and objects in which these new academic reading practices play a role. Critical relational terms used to describe and discuss these contingencies are virtual and actual.

### **Ontology of the Virtual**

The word virtual is not being used here to designate a thing as being unreal vs. real. Virtual is rather used in relation to something being potential versus actual (Delueze [1968] 2004, p. 263). Virtual libraries are evoked in library catalogs, records in reference management software<sup>2</sup>, or a simple list of references. These all represent virtual libraries as

---

<sup>1</sup> Docear changed its name from Sciplore Mind Mapping in Oct 2011; during my work on this thesis. For the purposes of clarity I will refer to the software as Docear in all the text following this footnote. All references to Docear Mind mapping in interviews or elsewhere have been replaced with the name Docear. Docear was officially rolled out in February 2012.

<sup>2</sup> At the time of writing the popular reference management software programs are Endnote, Refman, Mendeley, and Zotero. They are designed for the personal management of bibliographic data and are a

they reference only the potential to exist as actual discrete collections. It is in the post-Gutenberg age of multiple copies that such virtual libraries gain credence. This is because bibliographic information mostly corresponds to a run of copies, rather than a particular publication. This means an individual has a greater chance of creating an actual personal library out of a virtual library evoked, for example, by a bibliography. The reality of multiple copies thus enhances virtual libraries represented via bibliographic citation conventions.

Actual libraries, on the other hand, consist of publications already acquired and collected on your shelf or in a folder of PDF files on your computer. Actual libraries are often less varied and inclusive than their virtual counterparts by the necessity of what is their main advantage—that they exist for the user. Colloquially, the terms digital and virtual are often used synonymously. I hope it is clear from the start that in the ontology of this research virtual is not synonymous with digital.

### **Findings**

This research reveals an arena where digital text is actual— present, tangible, and workable. This arena involves new technologically-assisted practices related to reading academic texts in electronic format. Maintaining folders of PDF-formatted academic text files forms the basis of these new practices. The files in these personal libraries do not include mere references or links out to other print or electronic text in other locations. They contain the actual digital texts. Users are able to annotate, fragment, link to, and structure texts in relation to one another more freely than ever before. The language of every interviewee in relation to their work with these digital facilities is concrete and tangible. At

---

common element of academic reading and writing practice; often used to track and cite traditional forms such as books, articles, reports, theses etc. However, they are flexible systems and can be used for many other types of research information management.

the same time they make associations to an enhanced understanding of the content of a particular text or, more often, a collection of texts.

So what is it about collections of digital texts maintained by readers in a standardized, full-text format? There are many reasons to be critically aware of the effects of standardization in the building of library collections, large or small. “Collections regardless of format” has been a kind of mantra for librarians since I have been one. However, this study concludes that credence should be given to how a standard file format like PDF, when combined with mind mapping functionalities like Docear’s, can rework how one goes about understanding a text and relating it to other texts. The interviewees in this study all describe the new practices as marking a greater scholarly command over discrete bodies of academic literature. Such reading practices (consumption of text) warrant attention from individuals and organizations involved in the production and use of academic text.

To summarize my approach in this conclusion I turn to Madeleine Akrich in her ANT study of technical objects:

The problem is not one of deciding whether a technology should be seen as an instrument of progress or a new method of subjugating people. It is rather to find a way of studying the conditions and mechanisms under which the relations that define both our society and our knowledge of that society are susceptible to partial reconstruction. To do this we have to move constantly between the technical and the social. (1992, 206)

This does not mean that concerns about subjugation and progress are not valid. She is rather arguing that objects of technology are part of the dynamic “infrastructure of relations” (Star, 1999; Bowker 1994) between humans and objects. The goal of this work is not to identify static objects of scorn or praise. Instead it is “studying the conditions and mechanisms under which the relations that define” academic text are “susceptible to partial reconstruction.” The goal is to create a perspective from which to better gauge how to proceed with work related to academic text.

## 2 Intentions and Context

Reading practices tend to be solitary experiences that entail a direct relationship between human individuals and the technologies of text. This is what makes them difficult to reveal ethnographically (Hine 2000, 54). It can also make them a less socially-fettered reflector of interplay between individuals and machines. Docear turns out to be a very personal tool, in the analysis I come to call its users “lone users.” So how does the solitary nature of its use relate to a field that is known for revealing the social forces in scientific knowledge and technological innovation? It offers more than a challenge of methods. It is also a challenge of framing.

This two-fold framing challenge relates to approaching individual practice related to formal knowledge in academic publications. First part of this challenge is to look at an instance where direct social/organizational activity is especially scarce in the micro. Second is a challenge of inserting practice into a frame seen to embody its antithesis— formal knowledge. The following quotation comes from an introduction by ANT practitioners to an anthology about knowledge practice in organizations (The volume itself is not limited to ANT or STS based articles.)

We must develop new linguistic and conceptual repertoires about knowledge that free us from prevailing notions that depict it as the static result of the thinking of disinterested and autonomous individuals and that suggest that knowledge can be stored, transmitted, and circulated to other individuals able to assimilate it into some form of mental or material repository. (Nicolini, Gherardi, Yanow 2000,7)

This study sets out to work through blanket statements like this one; to break out bits for better observation. While the quotation’s call for a vocabulary of practice is a useful and important shift, it is also blinding to what is practice within the reading of academic text. In response I will rework of the abovementioned quote in a way that attempts to reinsert a less

“static”, “disinterested” perspective on the practices of consuming formal knowledge. It is a definition of academic reading practice that functions for this study. In addition it incorporates elements (given in quotations) that I want recover from the above Nicolini, Gherardi, and Yanow critique.

Academic reading practices relate to the often “autonomous activity” of reading whereby some combination of “mental and material repositories” are used to “assimilate” aspects of “knowledge [that can be] stored, transmitted, and circulated” in some format via primarily text-based academic publications.

This study does not question the thinking that practical (tacit) knowledge is a precursor to reflexive theoretical knowledge (Polanyi 1958). It is rather looking closely at reading practices that surround that particular reflexive, which has been an undeniable component to academic and scientific inquiry thus far. Reading academic text embodies the many synapses that make possible the putting together of Polanyi’s “jigsaw puzzle of science” (Polanyi 1962). This is a metaphor through which he calls for a belief in the crucial value of scientists “freely making their own choice of problems and pursuing them in the light of their own personal judgment.” It is a metaphor where scientists independently contribute to an emerging picture of reality. It is “coordination through mutual adjustment of independent initiatives – of initiatives which are coordinated because each takes into account all the other initiatives operating within the same system”(1962, p.2.). This idea of scientists’ independent action has since been successfully battered on various fronts by the field of STS (Collins & Pinch, 1998; Latour & Woolgar, ([1979] 1986). As successful metaphors often do, this one makes blurry the graininess of the reality it is trying to represent. The social/economic forces placed on scientific inquiry are lost as well as the role played by the mediums of academic communication. The idea of puzzle pieces simplifies the actual complex physical pieces of collective scientific enquiry—academic literature. This study looks at the latter.

## 2.1 Choice of Subject

My interest in pursuing this sociological study of text technologies has its roots in a curiosity about my own professional focus on the content of text. I have been an academic librarian for 20 years. During this time a landscape of networked digital equipment and its variety of electronic platforms and formats for content have fully emerged. To do my work I needed to understand how to function in this landscape; but for a long time I defined my job as being about content—the content and its users. Formats and mediums were a practical sideline related to services. Frustrations arose when I felt that the nature of my institution’s focus on emerging information communication technology overshadowed the content.

The director of a large academic library where I worked had the following motto on his email signature that he encouraged us all to use as a motto for the library:

“We put the information in information technology!”

A small stir was created by the spreading of a cynical version of this motto:

“We put the technology between you and the information!”

I reveal this embarrassing chapter in librarian humor, because it did actually create a stir and it was a bit funny for us at the time. There was indeed a general professional fixation on content and its ultimate separation from technology. This fixation is surely not limited to librarianship but librarians are in a privileged position for observing the grainy elements that surround the use of academic text as a physical object. Librarians have slogged through the stuff of OPACs, CDROM towers, Dial-up Services, Internet, not to mention paid the bills. Through it all they have staffed the desks that fielded the actual questions situated in the interface between content and its technology.

Yet still the perception of a separation between content and technology has lead librarianship to at times treat technology as a neutral hurdle. This is illustrated by the policy recommendations in the 1990s for shifting from the idea of building collections to the idea of managing content. It was a shift to a “vocabulary that focuses on content and transcends the very static dichotomy of ‘ownership and access” (Budd & Harloe, 1997). Here content is seen as independent; it transcends the mediums through which it is transmitted. Seeing ownership and access as a static dichotomy is a belief undermined by this study. Can a focus on content have made many of us unable to properly see the emerging mediums? This possibility was revealed to me by 1960’s media guru, Marshall McLuhan, who said about brain surgery and night baseball games that they

. . . are in some way the content of the electric light, since they could not exist without the electric light. This fact merely underlines the point that the medium is the message because it is the medium that shapes and controls the scale and form of human association and action. The content or uses of such media are as diverse as they are ineffectual in shaping the form of human association. Indeed, it is only too typical that the ‘content’ of any medium blinds us to the character of the medium. (1969, p.8)

This quote is from the book *Understanding Media*. Its macro perspective entails a sweeping historical survey of technologies as extensions of the human body. It catalyzed my enrollment in the ESST program that now has me delving into a micro environment where humans and text technology interact.

In my review of the literature for this study, I learned that there is little research on academic reading practices. In Christine L. Borgman’s (2007) award winning<sup>3</sup> and thorough 265 page analysis of *The State of Digital Scholarship* there is little mention of scholarly reading behavior even under the category of “seeking and using information,” which she describes as “among the most researched topics of scholarly practices, with thousands of

---

<sup>3</sup> Borgman’s book won the "Best Information Science Book" by the American Society for Information Science and Technology (ASIS&T) in 2008.

studies since the 1940s” (2007, p. 155). The survey of research she presents that relates to reading, as opposed to searching, is limited to quantitative findings such as numbers of articles read and publication dates of articles read and citations to articles. She states that research in the area of reading is based primarily on bibliometric indicators instead of direct observation or interviews (p. 157). Christine Hine (2000, p. 54) attributes this same lack to the problems of making the interpretations of texts ethnographically visible. “Producers [writers] embody their concerns in the technologies they produce, and the work of constructing a technology is highly visible and observable. Users [readers] leave no visible marks on technologies, and interpreting the technology is often something they simply get on and do” (p.54).

Considering the resources put into scientific research, its writing up, publishing, and dissemination it is curious that there is not more ethnographic research that looks at the situated consumption of this enormous investment. It is curious that inquiry stops at downloads and citation counts and looks no further. I am considering the degree to which this is a question of framing? In discussions of science’s role in public controversy Brian Wynne articulates the role of framing within the STS subfield of the Sociology of Scientific Knowledge (SSK) he raises concerns about how propositional knowledge-claims “harbor and protect tacit commitments to particular meanings or salient questions, which if just presumed and left implicit can become effectively imposed with no collective negotiation” (2002, p. 403). Such tacit commitments are hidden in the framing of any issue for any number of reasons from to politics or personal comfort. Framing is always required to see an issue

while at the same time it always limits what is observed. Revelations in this regard often boil down to what is left out of an inquiry's frame. With this in mind I offer a quote from Lars J. Lundgren in the article titled *The Utilization of Research*. [Highlighting is the authors.]

The first requirement for encouraging the utilization of a research result is that it exist in the material world, which normally means via a journal or other publication. This is easy to verify: either the result is available at a given time or it is not. Next, it must be disseminated in various ways so that it reaches those who can be expected to want to learn about it, and who may need it for making decisions, or for conducting an enquiry. **They have to become familiar with it in some manner.** A researcher always has an interest in getting his/her results published, primarily in scientific journals. Even if the results are relevant to practice, though, it is not certain that a practitioner will succeed in locating them. The researcher, or someone else, must then **realize their relevance** and see that they are also spread in other ways. (Lundgren, 2000, p. 153)

The highlighted excerpts assume that someone at some point actually read and interpreted the text in question. They are perhaps the vaguest, most passive, least concrete statements in the paragraph. They are akin to Hine's statement about reading being something scholars "simply get on and do." When Lundgren says "becoming familiar," it is not clear if he means actually reading a publication resulting from research or finding it in a bibliography. However, he calls text the "normal" means by which research "exists in the material world" That it takes the form of a "journal or other publication" says relatively little today about the text's actual physical medium or infrastructural context. Text content is considered alone in this paragraph without a clear medium or clear situated use. The practice of reading is absent.

The bits of text analysis used so far are not a critique of the scholarship or objectives of these authors. They are rather a way of recognizing a dark area in the micro arenas of scientific inquiry; arenas that have otherwise been quite well illuminated by the field of STS. ANT, for example, has shown that practices related to the production of scientific text are critical to our understanding of scientific inquiry. This was catalyzed by Bruno Latour and Steven Woolgar's seminal work *Laboratory Life: The Construction of Scientific Facts* (1986).

The consumption of scientific text, however, has yet to be illuminated. That is what this study is attempting to do. It does not let statements like “simply get on and do” or “become familiar with in some way” sail past. Using ANT this study employs digitalization as a point of change in relations between humans and text objects to reveal the performance of scientific literature as it is being reconstructed.

## **2.1 Notes on My Approach to Writing Ethnography**

This section arguably belongs in methods but I would rather have an explanation of the manner of writing come earlier on in the reading. Ethnography is a main informant for the branches of STS that this research is based on. This is often a multi-sited ethnography and one that has distanced itself from traditional ethnography’s insistence upon a sealed-off culture and an objective observer (Marcus, 1995). There is a precedent in STS, and ANT in particular, for established practitioners in a field to use its perspective and methods as a reflexive lens. I have entered the realm of this type of ethnography, which further challenges traditional ethnography’s standards of objectivity. Though not in libraries or among fellow librarians, my research was conducted within the mediums and networks that define the key issues of my current profession as an academic librarian. In addition, during the course of my study I have been using the tool (Docear) that is the object of this study as an aid in the writing of this very document. My familiarity with it becomes a tool for both methodology and analysis since a shared understanding of the software with the interviewees made sense of their remote descriptions of situated use.

I rise to these challenges to objectivity by not attempting to rely on it too strictly. Trying to simply “fix a world on paper,” according to Latour (2005, p. 121), will not function because the mediating constraints of writing are not taken into account. Reflexivity is

beneficial when it enables awareness that text is a constructive act rather than a straight forward reflection of reality (Denzin, 2010). A few small indulgencies may carry the reader along but the main intention with the personal narrative elements in this thesis have been to make clear how my experience informs my intentions in the framing and interpreting of both subject and object. Inspired by William James' radical empiricism, I am not seeking absolute truths but rather studying local realities with the goal of informing a work-a-day life, so that "we may know what is in the wind for us and get ready to react in time" (1912, p. 96).

And lastly, I will note that at a few points I use Actor Network Theory's manner of giving voice to non-human objects as a means of making clear how their obdurate qualities are, in a sense, negotiating. Among the introductory quotes is one by Steven Woolgar (1991, p. 61), "Machines as text." I agree that machines, by embodying a dialog between design and use, necessarily carry messages. Speaking for them can be seen as an appropriate way of reading them. I mostly have read them through use, observation, and inquiry. Woolgar encourages one to "explore the metaphor" and I accept his invitation.

## **2.2 Choosing an Object**

So following that defense of reflexivity, I will transition into a description of how I chose Docear as an object of study. I was looking to learn about emerging text objects and individuals' use of those objects— some way to make visible what was becoming new practice regarding acts of reading academic text. Something where the user left "marks" as Hine (2000, 54) put it. Through my work as an academic librarian I had encountered the following two software programs, which become the possible candidates. Describing them here illuminates the basis of my choice.

## **1. Annotation Tool**

This is a software program developed and used in the Netherlands. I found out about the software during a presentation made by a professor of education at the University of Utrecht. It creates a collaborative space accessible exclusively to participants in a course where they can collectively annotate digital documents saved in a common networked space. It is a simple construction. The document text is made available on the right in the larger of two frames. To the left are the comments of other users/readers linked to specific points in the text. When logged on you can make comments and edit your own comments and see the comments of others. You can also comment on others comments. The professor mentioned that they had structured use such that students were required to add comments before the reading was scheduled to be discussed in class. The concept and software was a custom project operating out of his particular faculty of education.

## **2. Docear**

I found Docear in the same way that all of the users who I eventually interviewed found it. I was surfing the web for something about mind mapping and academic work. And like all those I interviewed I do not remember exactly how I was searching or why. Once found, a promotional video provided a clear presentation of its functionalities. It showed how the system enabled users to dynamically link academic publications in PDF format to digital mind maps. It facilitated digital mind mapping as a way to organize a personal archive of academic literature together with online personal annotations to those documents. It is a combination or mash-up of mind mapping and PDF reading tools.

I had decided to focus on software because of its relevance to today's practice in a digital environment. I chose to focus on the use of one software as a strategy whereby discussions of academic reading practices would be situated in one common digital arena; a situation that both interviewee and interviewer could discuss concretely. Software as a relatively fixed, standardized object meant that we would be immediately situated in a way that had common discussable reference points. It was only after I completed the interviews that I made the connection between these common discussable reference points and Lucy Suchman, Randall Trigg, and Jeanette Blomberg's (2002) work with situated use and prototypes that I will discuss in more detail later.

Annotation Tool and Docear both facilitate the reading of academic text. Both enable the reader to leave visible marks as they read. While I was not entirely sure, I assumed that, unlike Annotation Tool with its classroom application; Docear was a product people worked with independently. (This proved to be true.) I thought Docear would get me closer to the less accessible, direct human/text interaction I was trying to root out. Not to mention the fact that it was something that I too could independently come to a working understanding of. I was engaged in the processes it was design to assist— writing an academic work. And, as mentioned above, I had already decided to err on the side of reflexivity over objectivity.

In the end my ultimate choice of Docear was due the immediate support of one of its developers and co-founders, Jöran Beel. He offered to post a call for interviewees on the official Docear blog and within days I had 10 potential interviewees— a windfall for a student in an accelerated Master's program. I chose Docear.

## 2.3 Description of Docear

Docear, is a non-profit software program founded by two German PhD students and based out of the Otto-von-Guericke University, Magdeburg Germany with elements of the project based out of the University of California, Berkeley USA. At the time of writing its founders Jöran Beel and Béla Gipp were using Docear's development as a basis for their PhD research. The project had received a year's funding from both the European Union and German Federal Ministry of Economics and Technology in July of 2011. Its main method of marketing was through academic channels and via its website.

As of December 2011 its website promoted Docear as:

. . . an 'academic literature suite'. It provides to researchers what an office suite provides to office workers. By that we mean, Docear supports you in doing all the tasks related to literature management.

1. Finding new literature (searching and exploring)
2. Organizing literature (reading, annotating, and retrieving)
3. Creating your own literature (drafting, writing, referencing, and publishing)

This description is based on the plans for a rollout of the final alpha version, which took place in February 2012. In the beta version only the features related to "organizing literature" and what I am calling academic reading practices were fully developed.

### **Docear's Statement of Academic Intent**

The focus of this research is Docear the software program as an innovative object for the five users interviewed. I am not studying the Docear team members or Docear as an organization. However, a presentation of Docear's statement of academic intent is useful as background

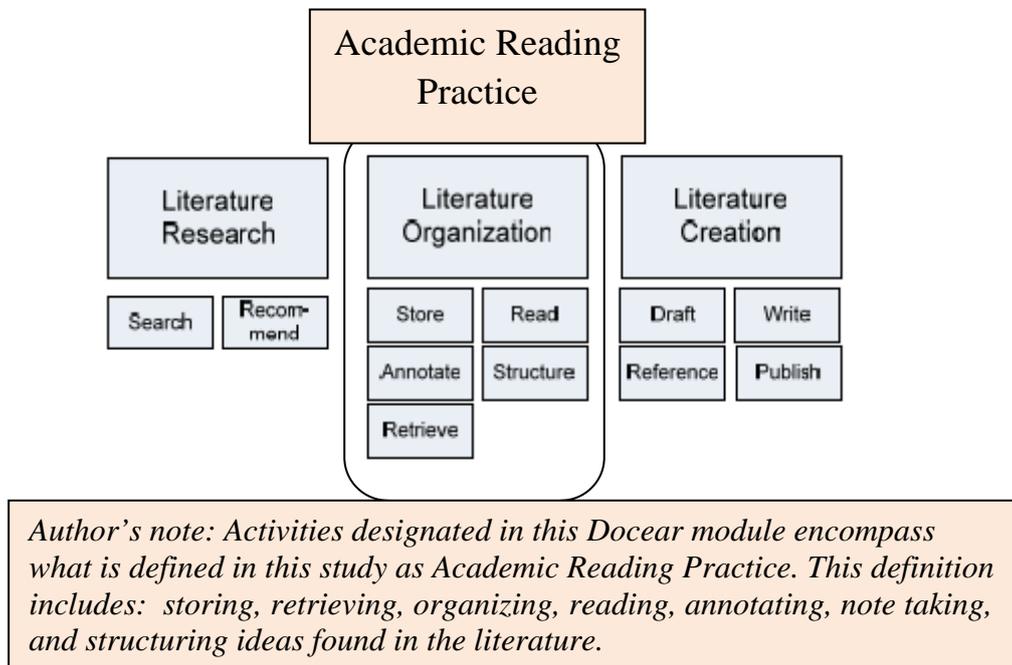
information. After receiving a year's grant from the German Government for 100,000 euro the following statement was published by Beel on their blog:

. . . we are extremely happy to have this scholarship as there are no strings attached. It's not a loan we would have to pay back, and it's not an investment from some investors expecting a huge return on investment. . . . After the year, we hope we will be able to continue the development of Docear in a similar way as Zotero is doing it, with research grants and funding from non-commercial organizations.

I conducted an in-depth interview with the co-founder, Jöran Beel. In it he clearly states that his intentions for the project are academic in nature and based on his applied research in user modeling.

### **Docear's Functionalities**

An overview of the planned functionalities for the entire suite is presented in figure 1. The blue boxes are a copy of a graphic created by Beel (Beel et al 2011, 1) and his colleagues. I have added brackets and notes that identify academic reading practices as they are defined for this research. Shown are the different processes Docear plans to support and how these relate to functionalities in the different modules of the suite. I will not be addressing the "Literature Creation" aspect with its word processing, reference management and publishing software. Neither, will I be addressing the "Literature Research" element, for which there is being developed an open searchable database of academic literature called Mr. DLib. Mr. DLib is an academic search motor with similarities to Google Scholar, but not presented as a competitor. The "Literature Research" element uses Mr. DLib as the basis for a recommendation service for academic literature. The service works off of content in an individual user's Docear mind maps.



(Beel et al 2011, p.1)

**Figure 1: Docear's Presentation of their Academic Literature Suite**

(The brackets and the Academic Reading Practice labels are added by the author to show how these two nomenclatures correspond with each other.)

The focus of this research is the mind mapping functions in Docear's "Literature Organization" module, which was well-developed at the time of the interviews. Docear links mind maps to a users' library of PDF files to create a dynamic means of organizing literature. How this is actually performed will be described in detail in the analysis section. I am including a brief working description of mind maps and PDF files as they relate to the situated use that forms the basis of the empirical research for this study. Included in this is a brief discussion of the role writing plays in what is a study of academic reading.

## Defining Mind Maps

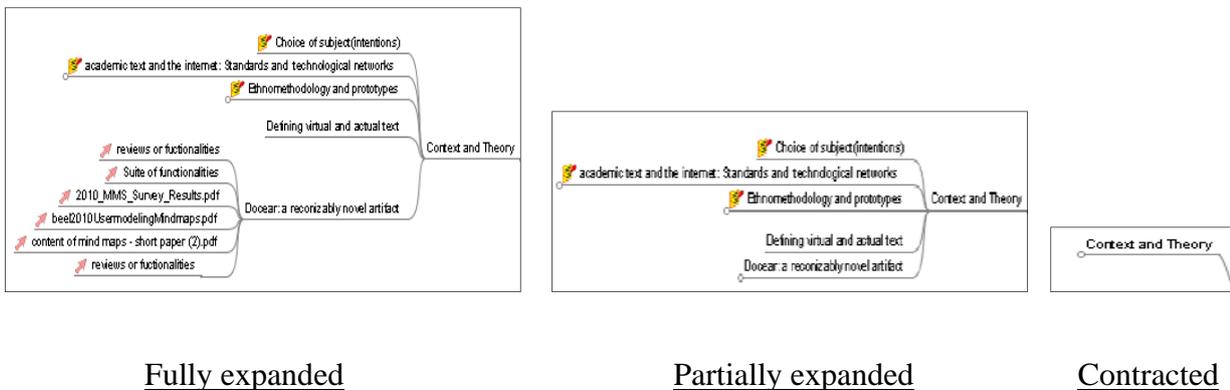
Discussions and definitions of mind maps can be based on various premises. Generally, mind mapping can be described as a way of visualizing concepts. There are terms that predate the term mind map, such as the more academic term “concept maps” or the more colloquial British term “spider maps.” Visualizing mental processes can be done through a wide variety of graphic representations—connecting nodes, concentric circles, stacked boxes. Theoretical and methodological discussion can be based on how particular fields might use mind mapping or concept maps for example in education or psychology. (Wheeldon and Flaubert, 2009).

This research defines mind maps out from specific physical and structural characteristics, particularly in relation to how they have been translated into a digital environment. I am using the definition of mind maps given by Jöran Beel and his colleague at Docear, Stephan Langer (2011, p.1). They describe the following varied practical uses, “brainstorming, note taking, document drafting, project planning and other tasks that require hierarchical structuring of information.” Then they go on to describe an example of their quintessential mind map:

. . . it has a central node (the root) which represents the main topic the mind map is about. From this root node, child-nodes are branching describing sub-topics. Each node may contain an arbitrary number of words. This way, a mind map is comparable to an outline but with stronger focus on the graphical representation. Mind maps created on a computer may also contain links to files and hyperlinks to websites, pictures, and notes. (Beel and Langer 2011, 1)

I would like to add a description of the ability to contract the above mentioned nodes. This means that the mother node to a complex nesting of child-nodes and branching sub-topics can be contracted such that only the mother node and a link to expand it again are visible to the

user. I have heard this called expanding and contracting, opening and closing, and folding and unfolding, see figure 2 below.



**Figure 2. Expanding and Contracting Mind maps.**  
 (Screen shots of one of the author’s Docear mind maps)

The expand-and-contract feature allows for enormous amounts of information to be immediately accessible on one screen. This sheer volume can be seen more clearly in screen shots in figure 6 .

### PDF Software and Academic Publishing

What makes Docear’s digital mind maps new is the way in which they link to PDF files, simultaneously importing bibliographic information and bookmarks made in PDF viewers. The Docear team do not describe or define PDF software in their product materials or articles. It is not useful to go into technical details of PDF software her. Instead I offer a description of PDF software that is relevant to academic reading practices. It is common to use PDF files for academic digital publishing. PDF stands for Portable Document Format. While it was developed by the Adobe Company the actual PDF file format is now an open standard for document exchange. This is reflected in the proliferation of PDF viewers for which Docear has made itself compatible. In a PDF the original academic text content is protected from alterations. Unlike a word processing document, in PDF you cannot go in and cut and paste

and move around and erase elements of the original text without changing the integrity of the format. While it allows for hyper linking it is otherwise a fixed or flat image.

When an academic article is published on the Internet in PDF format the original academic text is therefore relatively stable. For purposes of academic literary conventions this is critical to the PDF's usefulness since the authority of the scholarly process relies on being able to return to an intact version of the original text. At the same time that the PDF document is intact, most PDF viewer software supports annotations in much the same way that one works with a printed book: by highlighting, adding marginalia, and bookmarking. Marks are left without changing the integrity of the original work.

### **Reading to Write**

I have intentionally not included the word writing in my classification of academic reading practices. However, forms of original writing are integral to reading practices in Docear. Writing is involved in annotating PDFs and writing has to exist at a minimum in order to label the structural elements (nodes) of a mind map. In addition, text boxes can be used within Docear mind maps for notes and partial drafting of an eventual publication.

The lines between consumption (reading) and production (writing) are arguably less defined with academic literature than other forms such as journalism or fiction. Writing academic literature most often requires extensive reading of academic literature in the form of a literature review, mentioned by all of those interviewed. So, one is often writing it as one reads it. At the time of writing, Docear software required that work done in mind maps be manually copied over to a word processor to be made publishing-ready by most academic standards. I am focusing on writing that took place in Docear mind maps and am calling it note taking and relating it to a process of understanding existing text.

# 3 Micro Environments, Infrastructure, and Situated Use (Theory)

## 3.1 The Micro & Infrastructure in Actor Network Theory

In this chapter I will be filling in elements of the theoretical basis for this research.

Throughout this work I will be drawing from the work of both founders and practitioners of Actor Network Theory (ANT), Bruno Latour, Michel Callon, Madeline Akrich, Steve Woolgar, Lucy Suchman, Susan Star, and Geof Bowker. They have developed an approach whereby one understands and studies a world from the perspective of dynamic networks called Actor Networks. There is often confusion related to their use of the word network. In ANT a network is dynamic, not fixed like one thinks of a subway network or a computer network. ANT networks are made up of actors or actants that are not necessarily human. They are also called a hybrid networks. Actors and actants act and are in continuous change in relation to one another, which creates the dynamic.

Underlying (ANT) is built the idea that the use of the word social, as it is often used in social science, reflects an unhelpful way of looking at the world. This critique of the word social is made particularly clear by Bruno Latour when he discusses its use as an adjective to describe things. “problems arise . . . when ‘social’ begins to mean a type of material, is if the adjective was roughly comparable to other terms like ‘wooden’, ‘steely’, ‘biological’, ‘economical’, ‘metal’, ‘organizational’ . . . “(2005, p.1) Later in the same text he comes with examples: “social norms”, “social customs”, “social structure” (p. 67) An Actor Network includes both humans and nonhumans in situations that are undergoing constant negotiation

or “process of assembling.” Using the word social to separate out human interactive forces as something fixed or a “specific type of ingredient” is the main ANT critique of its use.

Within ANT objects themselves are said to negotiate, a classic example of this is when Michel Callon (2007, 69) describes scallops “negotiating” with researchers and becoming “dissidents” of a project in which they are being farmed. ANT is criticized for using this form of anthropomorphizing in place of “scientific credentials” to analyze the actors involved, especially in the case of natural phenomenon like scallops (Collins 1992, p. 316). I will include here a quote from Latour that I think shows these linguistic techniques to have useful analytical applications that do not replace a scientific authority. It is useful as it relates to the research value of something new and technical. It also represents a “reassembling of the social,” which is the title of the book.

The first solution is to study innovations in the artisan’s workshop, the engineer’s design department, the scientist’s laboratory, the marketer’s trial panels, the user’s home, and the many social-technical controversies. In these sites objects live a clearly multiple and complex life through meetings, plans sketches, regulations, and trials. Here, they appear fully mixed with other more traditional social agencies. It is only once in place that they disappear from view. This is why the study of innovations and controversies has been one of the first privileged places where objects can be maintained longer as visible, distributed, accounted mediators before becoming invisible, asocial intermediaries. (Latour 2005, 80)

This is a good description of how objects are social and can be involved in a meeting or a plan. Madeline Akrich (1992, p.206) uses the term “obduracy” to evoke the way in which objects are tough negotiators that cannot be overshadowed by something described as “social” force. It is from this theoretical standpoint this research is looking to the hybrid network in which academic text resides. This includes both its physical mediums, and the contextualized practices of human readers within it.

Again, referring to the above quote, it is at the micro level that ANT works as a way of seeing. One must work at the level where the networks are continuously recreated since there one can see the relations in the Actor Network. The subject for this research, the practice of reading academic text, is not controversial and it is very much in place. Instead of looking to controversy as a revealing negotiation point this project looks to innovation; that is innovation the noun. The focus of this research is not how one innovates so much as seeing what is happening around an innovation; what is disturbed and remade. If we return to Latour's list above of places where "objects live a clearly multiple and complex life," Docear falls into the category of "trials." A trial is not exactly what the object of this study is. At the time of this study, Docear was the beta version of a software program, but it has the same qualities of being a new thing that is still working out the kinks. It is also akin to a prototype, which is a form taken up in the next section on ethnomethodology. Its kinks make it visible as a mediator. In this case things revealed include relations between reference management software, actual & virtual publications, time pressure around a publishing deadline, the size of a computer screen, and the act of reading. These are examples of what comes into play while understanding the wider infrastructure.

Susan Leigh Star offers a definition of infrastructure that she in turn rejects as not useful to understanding infrastructure: "Infrastructure as a system of substrates—railroad lines, pipes and plumbing, electrical power plants, and wires. It is by definition invisible" (Star, 1999, p.380). It becomes a black box or "intermediary" to use Latour's term above. It simply disappears from view and is not helpful to understanding. In Star's project with Karen Ruhleder (1994, p253) they argue instead for a definition where the central element is relations as opposed to fixed things. They further develop this idea through a methodological process called "infrastructural inversion" borrowed from Geof Bowker (1994), that they say;

de-emphasizes things or people as the only causes of change, and focuses on infrastructural relations (e.g. between railroads, timetables, and management structures in bureaucracies). It inverts traditional historical explanations and reveals how choices and politics embedded in such systems become articulated components. Substrate becomes substance. (p 253)

Their discussion is related to a project to develop a custom research software program on a large scale. It is the basis of their argument to change systems development and deployment that is,

based on a set of rationalistic or “mechanistic” ideas about artifacts and infrastructure, [which they say] assume that tasks to be automated are well-structured, the domain well-understood, and that system requirements can be determined by formal, a priori needs-assessment. (p 253)

The subject of this thesis similarly looks at research software, albeit one with differing relations in that it does not work to unify an organization. It is not collaborative in the same way that Star and Ruhleder’s research object is. However, this does not negate the benefits to this study of attempting a methodology that takes its starting point from the principle of infrastructure as relation, and that avoids “formal, a priori needs-assessments.” There is much in the relational infrastructure of academic text that can distract one into a technical and, therefore, less revealing view. If one does not fixate on the content it is easy to fixate on the technical.

## **3.2 Ethnomethodology: Prototypes and Situated Use**

Lucy Suchman, like Star, Bowker and Luhleder, also criticizes the more mechanistic approach to systems development and deployment. In her work with artificial intelligence Suchman specifically looks at and undermines the traditional take on planning, specifically how it makes use of other’s plans; what Star and Ruhleder have called “a priori needs-assessment.” Plans for Suchman are one critical step removed from the reality of the relations

of an infrastructure or local situation. In her book *Plans and Situated Actions* (1987, p. 27) she observes how users interact with machines and how that relates to their plans and those of the machine's creator. Critical here is how the creator may have worked off of the potential user's plans, which translates easily to perceived user needs. Suchman rejects the traditional concept of assessing user needs as a way of understanding potential use. If you ask a focus group how they will use a potential product, they will give you a plan and you can be quite sure, according to Suchman, that they will not follow that plan.

This is based on Suchman's take on our how we understand the users at all. Here she looks to elements of Harold Garfinkel's theory of ethnomethodology. According to Garfinkel a hidden reality of human life is that it is nearly impossible for us to understand intended meaning in basic human conversation for "purposes of conducting everyday affairs" (1967, p. 41). He places the blame on an engrained tendency for speakers to have unreasonable expectations of how the listener can or will set themselves in the situation.

The anticipation that persons will understand, the occasionality of expressions, the specific vagueness of references, the retrospective-prospective sense of a present occurrence, waiting for something later in order to see what was meant before, are sanctioned properties of common discourse. (Garfinkel, 1967, p.41)

For Garfinkel a conversation must be situated otherwise it is a constant grasping and groping and waiting and looking back in order to make sense of words that are otherwise lacking in sufficient context to be of much use. His work in ethnomethodology is about getting down to the micro level in much the same way as ANT. However, Garfinkel did not have the same approach to a hybrid network. His were social contexts. He uses the concept of indexicality (2002, 113) to draw out reference points in our social lives that allow us to understand each other. That context can go missing in everyday conversations. This is a complicated methodology that I have not carried out. I am rather using Suchman's application of it as an analytical tool.

Specifically, I am taking things from a practical, applied approach to ethnomethodology developed out of her way of using a prototype in a study done with Randall Trigg, and Jeanette Blomberg (2002). I approach Docear in many respects as their type of prototype. Prototypes provide Suchman, Trigg, and Blomberg with tendencies surrounding the descriptions of the situated use of technology. Their research is far more involved in the technology than mine. They are directly engaged in creating the prototype, which is a computer filing system combined with a classification scheme. They're not interested in the text to be filed or the system but rather in the procedural work of document coding. In the same way I am not so much focused on the technicalities of Docear or the subject matter of texts worked with there. I am rather looking at it as an example of how the things that make up academic text perform. In this we see how "prototyping simultaneously recovers and invents work requirements and technological possibilities."(Suchman, Trigg, Blomberg, 2002, p.166) It does this by creating something "recognizably new." It is this combination, in retrospect, that I was after in selecting an object to study.

A proper ethnomethodological study is dependent on situated use, and therefore requires that researchers are together with individuals as they interact with the object of the study. It is a detailed and technical undertaking that I have not performed in the course of this research. I will therefore avoid its highly technical language here. But there are two things I have taken away. One is the idea of situated use. I have not actually sat down with users and discussed their use while I watched them, which is what Suchman, Trigg, and Blomberg have done. Instead I have had conversations with users that were tied to their use of Docear, of which we had a common infrastructural understanding. The nature of the Internet and instantly shared replicated software programs provided the interviewees and me with a shared space. It is a space dependent on connection instead of distance (Hine, 2000, 104). For me

as interviewer the context of the interviewee's descriptions of using Docear can be confirmed as real and intelligibly familiar. We are both connected by it and it becomes a shared space.

When I call this a study of situated use it is specifically related to Suchman's use of this term and its basis on the "indexicality" of language (Suchman 1987, 58). An example of an indexical statement in Docear would be "how did you link that up to the mind map" or "can we expand that node?" or "will you link that to your mind map?" This project's interview circumstances did not allow me to follow user's actions in this way. However, my questioning was tied to specific obdurate software functionalities that made it possible to pull from the interview text descriptions that got as close as possible to "just what it is that we are up to, here and now" (Suchman, 1987, p. 203). This comes through in the analysis as I reveal individual's consistent descriptions of specific new activities as they become practice.

My research benefits from how Suchman has worked in this way with technical objects. She is not interested in the functionalities but in the performances that make up situated use. This is where we arrive at practice. A particularly useful prototype, according to Suchman, Trigg, & Blomberg, will be one that might be "accountable to extended networks of professional relations and concerns, both technological and otherwise" (2002, p. 166). They argue for looking closely at the situated use of new technologies that may have an impact on forms of practice that are critical to an extended network. Much in the analysis points to Docear doing just this.

Suchman's approach applies to practical work related to human machine interaction. She creates a technique out of prototypes that manages to simultaneously speed up and be present to the varied arenas of an innovation. Bruno Latour and Albena Yaneva describe such methods as similar to what slowing down the film of a bird flying did for the understanding of flight. ANT is instead trying to reveal things that are hidden not in speed but in slow and

dispersed processes like the building of a building (Latour, Yaneva 2008). Suchman and her colleagues' ethnomethodological technique do this by observing throughout the participation in an innovation. My research is similarly trying to reveal an innovation in text. It was not until the interviews were completed that I discovered Suchman, Trigg, and Bloomberg's technique for using prototypes was applicable to my analysis. I had entered into the situated uses of the beta version of a software program and observed ways in which reading practices were being reconstructed there.

### **Docear as a “Recognizably New” Object**

In the description of its “program details” on the website in December of 2011, Docear starts by asking the potential user:

Are you using mind mapping tools such as MindManager, FreeMind or XMind? And reference management tools such as JabRef, Endnote, or Zotero? And do you sometimes even create bookmark in PDFs? Then you should have a look.

This description evokes the potential user's existing situated activity to make Docear familiar at the same time that it offers something new. It brings together various existing functionalities designed to help researchers work with academic text. I had imagined finding an arena from which I could survey and compare the use of various relatively new functionalities for working with academic text. Instead, I ended up observing a brand new thing as it altered the relations between objects and people and lead to observable changes in practice. As these relations became clear, the degree to which they were partially created by the Internet became critical to my understanding. Using actual and virtual as relational concepts is critical to how I avoid a fixed view of the technical infrastructural elements of academic reading practices today.

# 4 Ontology of Actual & Virtual

Using Docear functionalities entails maintaining a digital library. The less tangible, often remote quality, of digital documents in an online setting has implications for how readers can interact with the text before them. Reading an Internet document online, direct from a website, is to read on a screen while the text document actually resides out on a drive someone else keeps. It is like reading behind glass. We cannot make it our own by rearranging it in relation to other texts. Nor can we make it our own by leaving our marks on it and bookmarks in it. It is virtual in a way closer to Hine's (2000, p. 65) description of the connotations of the virtual as disembodies or "not quite" real, which hints at the often synonymous use of virtual and digital that we are otherwise trying to avoid in this analysis.

As touched on in the introduction this thesis is defining virtual as something more akin to potential and definitely not unreal. More than a definition this is a structure that reveals the relations informing this study. Its source is the ontology created by philosopher Gilles Deleuze. Deleuze sees a kind of intertwined dichotomy of a virtual and an actual that are something like a potential and its actualization. He states that:

the only danger in all this is that the virtual could be confused with the possible. The possible is opposed to real; the process undergone by the possible is therefore a 'realization'. By contrast, the virtual is not opposed to the real; it possesses a full reality by itself. The process it undergoes is that of actualization. ([1968] 2004, p)

Deleuze's virtual is not a potential to be realized like a seed that becomes a tree and then no longer exists as a seed. His is a virtual that is a potential to be actualized in the manner of an idea that is the basis for action but continues to exist.

This ontological idea has been actualized within the field of librarianship by Jean-Claude Guédon in 1998 in a lecture titled "The Digital Library: An Oxymoron?" He was

speaking to a meeting of medical librarians. It was published in the *Bulletin of the Medical Library Association*. Here is how he introduces the idea of a virtual library:

The word “virtual,” as a phrase “virtual library,” refers to an order of issues altogether different from that of digitization. In particular, we must not confuse virtual with “unreal.” Despite appearances, virtual reality is not an oxymoron because virtual is opposed to actual, not to real. The virtual is nothing but potential and as such, it is reality (possibility) in the making. (Guédon, 1999, p.10)

Medical librarians were early practitioners of digital librarianship. Digital libraries here are contextualized for a professional community actually working with digital documents. By choosing Docear this research has distanced itself from libraries serving a public and focused on the personal library as its object. Libraries serving a public are designed around the potential use of many individuals not the actual personal use of one. Personal or private libraries, in terms of the history of the book, lead to today’s modern library through an “actualization process” (Guédon 1999, 10). This research chose to study the emergence of personal digital libraries as of means of revealing something of that same trajectory taking place in today’s changing reading practices.

An interplay between actual and virtual literature has been facilitated for centuries by the conventions of what is today called bibliographic metadata. This is neither an exclusively digital phenomenon, nor is it dependent on a technologically generated demi-reality that the phrase “virtual reality” often invokes. It is data that is separate but integral to a publication’s function—its reality. Bibliographic metadata makes, among other things, a citation in a text possible. A citation in a text is a reference to a publication that the author has drawn from, but which is not present to the citation. The citation evokes a virtual publication that the reader can either remember or acquire aided by the bibliographic data provided in the citation.

In this study I am discovering how the new practices in Docear reveal a shift in the digital arena between a focus on the performed benefits of the virtual to a focus on the

performed benefits of the actual. In other words the digital as an arena for evoking literature (virtual) begins to be replaced by the digital as an arena for keeping and using literature (actual). The performed benefits of the digital arena for academic reading practice has until recently been primarily about the virtual—citing, tracking, managing, and searching information about publications. Bibliographic data and bibliographic software facilities generate enormous activity on the Internet about potential or virtual reading. This study shows that in an academic context the actual, full-text, digital publications are often being printed out before reading, or the full text is purchased in a print version. In these instances, if the full-text digital publication is being kept at all, it is not with the intention to read from it. Observed reading practices in Docear will reveal a drastic break from this “infrastructure of relations.”



## 5 Methods

This ethnographic study is based primarily on five in-depth interviews with users of Docear during its beta phase in the first half of 2011. No weight is given in the study to specific qualities of the interviewees as statistically significant in defining Docear's market or user base in anyway. It was critical that all were engaged in research in an academic setting that entailed reading primarily current literature obtained electronically, specifically in PDF form. I discuss characteristics such as demographics, work environments, hardware usage, and subject areas as a means to better understand these individual's use. The intention in collecting this data and subsequent analysis is not to learn about the software's potential for developing its user base. It is rather to better understand micro environments emerging from the use of Docear's unique qualities; particularly as it relates to the practice of reading academic text.

Tasks related to reading PDF-formatted academic literature are facilitated by Docear and it is these features that were the focus of the research. My own personal use of Docear was critical to my ability to discuss its situated use in the interviews. I asked questions about whether academic literature not in PDF formats were incorporated into their research. However, I have not delved into possible limits to content due to Docear functionalities' compatibility issues. But this is certainly a worthy topic for further study. For example, the field in which I am doing research, STS, I could not rely heavily on a Docear facilitated library since so much of the critical literature is published in anthologies that are not available electronically. The broader implications of how Docear-type technologies might inadvertently act as filters, while I find important, was not a focus in methodology or analysis.

<b>Summary of Interviews</b> Author-conducted audio interviews took place May 2011 via Skype from Norway.	Assistant Prof. Communication +U.S citizen +43 years ♂	MBA student +Russian +26 years ♀	PhD student Intl. Business +U.S. citizen +44 years ♀	Researcher in applied biomedicine (PhD cand.) +Australian +47 years ♀	Researcher in Applied Nutrition (PhD cand.) +Italian +44 years ♀
Interview length	<b>1 hour 8 min</b>	<b>1 hour 38 min</b>	<b>58 min.</b>	<b>30 min</b>	<b>36 min</b>
Months since starting to use Docear (approx)	<b>4 months</b>	<b>4 months</b>	<b>4 months</b>	<b>4 months</b>	<b>18 months</b>
Place of residence at time of interview	<b>Texas, USA</b>	<b>Delft, Netherlands</b>	<b>Kuala Lumpur, Malaysia</b>	<b>Melbourne, Australia</b>	<b>Delft, Netherlands</b>
Degrees previously earned	<b>BA in Liberal Arts; MA in Mass Communications; PhD in Educational Research</b>	<b>BA computer securities; BA applied math and financial management</b>	<b>BA construction management; MA in computer integrated manufacturing</b>	<b>BA in Zoology and BioChem; MA in Protein Chemistry and enzymology; BA Philosophy; MA in Law.</b>	<b>MA Public Health and Nutrition; BA Food Science</b>
Hardware setup	<b>Mac laptop</b>	<b>PC laptop</b>	<b>Laptop with extra 22" monitor. Simultaneous screen use. Smart phone.</b>	<b>I phone, ipad, pc laptop</b>	<b>PC Laptop</b>
Remote drives (cloud locations)	<b>Dropbox; private &amp; academic accounts</b>	<b>Dropbox; private &amp; academic accounts</b>	<b>Dropbox; Academic account</b>	<b>Dropbox; Academic account</b>	<b>Dropbox; Academic account</b>
PDF Software	<b>Adobe</b>	<b>Foxit</b>	<b>Foxit</b>	<b>Nuance</b>	<b>Adobe</b>
Bibliographic Software	<b>Familiar with them but not currently using one.</b>	<b>BibTeX(entries unspecified)</b>	<b>Refman: (Total entries unspecified); 323 entries for items owned in print or electronic</b>	<b>Refman(entries unspecified)</b>	<b>Endnote (4500 entries)</b>
Approx number of academic publications in PDF format	<b>500 PDF files mostly articles.</b>	<b>unspecified</b>	<b>290 PDFs; 10 electronic books</b>	<b>300 pdfs</b>	<b>1000 PDFs</b>
Docear maps in active use	<b>1</b>	<b>4</b>	<b>2</b>	<b>5</b>	<b>3</b>
Access to full-text articles database	<b>Yes, via academic inst.</b>	<b>Yes, via academic inst.</b>	<b>Yes, via academic inst.</b>	<b>Yes, via academic inst.</b>	<b>Yes, via academic inst.</b>
Transcribed text describing daily use.  (This text was extracted here to show the degree to which the interviewees were established users of Docear. This interview text is not included in the analysis, which focuses on established user's descriptions of situated use.)	I could go days without opening it, even a couple of weeks when I was traveling to using pretty intensively in a single day. . . . When I'm doing the literature review, it's open even if I'm not actively doing something inside of it because that I'm reading and making this and highlighting the PDF.	Well, let's say, my morning starts at 7. At 9, I usually start working on my paper and that goes until approximately 6-7 PM. At 9, I start my TeX editor, my PDF viewer, my Docear and a couple of things like Skype. So let's say, I start it every morning for eight hours at least.	Extensively. It's probably the first application I turn on in the morning and the last one I turn off at night. [He had mentioned earlier in the interview that he was "going through his literature review right now."]	[He describes current focus on work in the laboratory.] . . . not doing that much reading . . . So in a working day at the moment, I can confidently say I'm not using Docear right now. But once my reading starts again, I [will] be back online.  [When asked if he would be using Docear, "quite a bit " during intensive reading periods he gave the following answer.]  Yeah, absolutely.	Doing the research is basically when you are doing, you use the Docear massively. Something like probably you have a Docear open 80-85% of your time while you are reading your literature. And then when you organize your paper structure or your chapter structure. And then probably you go down to probably to 50%, 40% when you're actually writing your chapter or your paper. So it really depends in which phase you are. And then you would answer either 95%, 90-95% or half of the time or less than that.

**Figure 3: Summary of Data about Interviewees (Context and Relevance to the Study)**

## 5.1 Interview Process

### Finding interviewees

At the end of May, 2011 I sent an email to the two founders of Docear and expressed an interest in using Docear as a case in my Master's research. I explained the project briefly mentioned an interest in individual or collaborative use of Docear. I gave the name of my department at the University of Oslo, The Center for Technology Information and Culture. I also gave a bit of my professional background as a librarian.

Jöran Beel replied to my email and was very positive. We had a brief telephone conversation. He offered to post an announcement on the Docear blog which was linked to its website, see appendix 3. The information in the call for interviews had a marketing and product development presentation that did not exactly represent my research intentions, but I assumed it would be how I conducted the interviews that would set the tone. Beel received the emails from potential interviewees directly and forwarded them on to me. I have no information about the contact he had with the ten users forwarded on to me between May and July of 2011. If there was some form for screening done it did not hinder me in locating five relevant interviewee objects.

Of the ten responses that Beel forwarded on to me I sent out eight standard emails regarding possible times, methods, and terms of agreement. Two of the ten from Beel responded too late for me to work into my schedule and I sent thank you emails instead. All of the ten users described themselves as engaged in academic research. Only one

worked independently of an institution of higher education, he did not reply to my direct mail.

The email I sent directly to eight of the respondents was standardized and included information about interview times and time zones, as well as terms of agreement regarding interview methods and data. It was suggested that they answer in the text of the email and reply. For a complete copy of the standard email see appendix 1. Of the eight emails sent, two did not respond and with one there were scheduling problems. That left me with 5 scheduled interviews with interviewees who had agreed to the conditions and agreed to being recorded. In addition, I conducted an in-depth, almost 100 minute interview with Jöran Beel.

### **Interview Technology**

All five interviews plus the interview with Jöran Beel took place in May of 2011 and were conducted using Skype audio. I decided that the video would be too distracting and would not add anything to my discussions of situated use. One of the interviewees offered to guide me through his work with Docear using a shared screen function in Skype. I was interested in this as a tool for remotely getting closer to situated use. However, since it was the third interview I decided changing methods and terms mid-stream would do more to confuse my results than improve them. I will strongly consider it for any future projects involving situated use of software.

The Skype technology was easy to use for the interviews and it was easy to add the plug-in for digital recording. Instability was a problem and two of the interviews were interrupted, but we were able to connect again and finish the interview. There were a few problems related to time changes. In order to overcome scheduling conflicts two of the

interviews took place in the middle of the night my time. Only one of the interviewees did not have access to Skype for the interview, but I was able to purchase Skype-to-phone minutes and call him on his mobile phone and still retain the recording functions.

For transcriptions of the interviews I used another Internet Service called Scribie. My main reason for doing this was to save transcription time on my end. And I figured since I had all of the audio files there was no danger of losing integrity by having a service do the first go through. A description of the transcription process is published on the internet, see appendix 5 for a copy. They mention that the transcribers will look up acronyms and proper names as part of the service. This could have been a problem since it is not a pure transcription, but the places I noticed did not have an impact on the analysis of situated use. There were instances with the one female interviewee where our voices were confused with each other, but this was easy for me to catch.

### **Conducting the Interviews**

The interviews were semi-structured and divided up into three sections. I prepared an interview script that I held fast to for the first context section. Then I was more flexible as regards the more open ended questions that followed related to situated use and perceptions of use. I tried to ask all of the questions and added a follow-up question if a comment seemed particularly illuminating of what I might be after. I asked in the beginning if they were willing to go over 30 minutes and three of the five were. This also had an impact on the degree of flexibility.

The first interview section included “context questions”, demographic questions and questions about education, field of study, and institutional affiliations. These were followed by questions about the users’ technology set up in terms of hardware, software, and support.

None of these questions were meant to gather information of any statistical significance but rather to be able to compare the users and understand their context from various angles—material, academic, and demographic. I moved through these questions quickly so as to have time to discuss situated use.

The next two sections were called “open questions” and “functionality questions.” Open questions were meant to draw out the story of their use. I ended with specific questions related to the various Docear functionalities. This order seemed counter-intuitive at first, but was recommended by my advisor and proved fruitful. Starting with the open-ended questions allowed me to better direct the specific functionality questions towards how they were actually working with the software. This came from my initial strategy of use the software as a shared set of functionalities that could serve as reference points in the interviews. It was in retrospect that I saw that is strategy’s interview results enabled an analysis of practice based on situated use.

## **5.2 Summary of the Interviewee Relevance**

This profile summary of the interviewees, confirms their relevance to the study. I have not done an in-depth summary of each individual user. I instead present them in the course of the analysis by theme, see figure 4: Summary of Interview Data.

### **Basic demographic information:**

The user interviews ranged from 30 minutes to 1.5 hours (approx.) At the time the five Docear users:

- Ranged in age from 26 -49.
- Represented 4 nationalities: U.S. (2), Russian, Australian, and Italian.
- Were located on 4 different continents: Asia, Europe (2), Australia, North America.

- Represented 3 fields of study: 2 students of management (MBA and PhD), 2 researchers in applied medical sciences (both PhD candidates), and 1 assistant professor of communications.
- Were 4 men and 1 woman.

Relatively speaking, for such a small sample it nicely represents different regions of the world and fields of study. All of the users were in their 40s apart from one who was 26. It's important to note here that we are talking about a networked digital arena where things are changing rapidly. Even though the age range of the users was less than 25 years there can still be considerable variation in what and how users recognize things or what represents change for them. This can be seen at times in the interviewee's responses but I have otherwise not addressed this variation.

### **Interviewees' Degree of Engagement in Academia**

All of the users were affiliated with an academic institution. One was an assistant professor in communications, two were established in applied natural science research as well as being PhD candidates, and two were students of Business Administration (one MBA and one PhD Candidate.) Both the range of fields and roles in academia served the research intentions. And all were established enough to be engaged in the technological network of academic text in terms of access to databases and standard academic ICT software and equipment. One can infer that this entails engagement in the broader infrastructure relevant to the common practice of reading academic text.

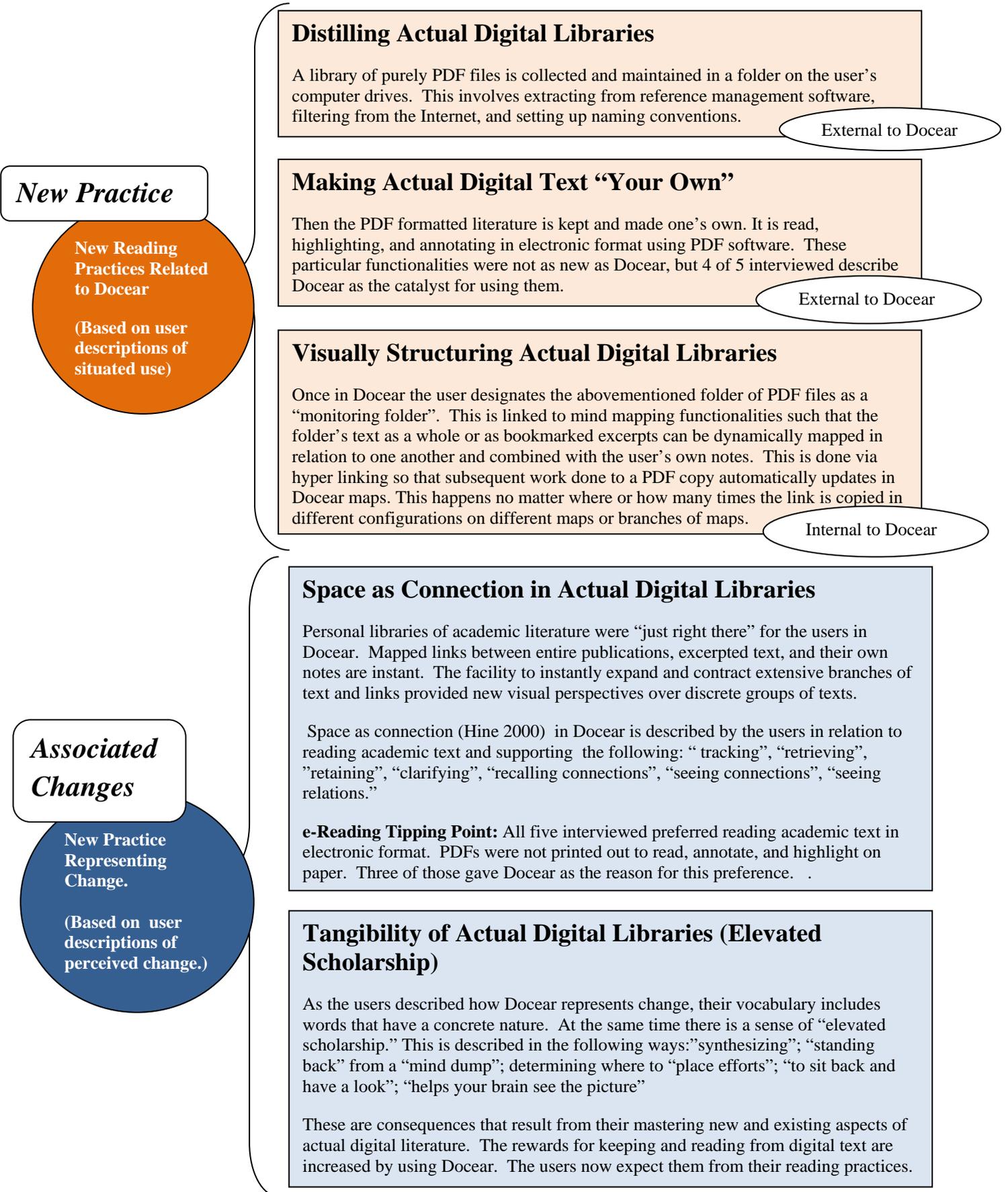
### **Interviewees' Degree of Engagement in Docear**

I am basing my research not on Docear features alone, but on how Docear's use influences research practices related to the consumption of academic text. It was critical to this project

that the users interviewed were not just interested in exploring the software as such. That would have told me something but it could not have been the basis of a Master's thesis about practice related to consuming academic text. It was therefore critical to confirm that they were engaged with the software to such a degree that it had the possibility to have an influence on their academic work.

I think it is clear from the analysis section that this is the case for all of them. A summary confirmation of this can be found in collected excerpts of text where the users described the scope of their use, this is found in Figure 3: *Summary of Data about Interviewees (Context and Relevance to the Study)* on page 32. Under "Transcribed text describing daily use" all describe a consistent use that is sporadic due to the fact that reading habits are periodic and related to intensive periods of literature review. That all of the users mention this gives further weight to my results related to descriptions of situated use. It is a further confirmation that for all of the users interviewed Docear has had an impact on sustained practices related to reading digital text. This is also made clear in the descriptions of situated use presented in the analysis section.





**Figure 4: Summary of Findings: Actualizing Digital Literature in Docear**  
(Situated Uses and Perceptions of Change)

# 6 Findings and Analysis

What sorts of human practice have begun to establish themselves around Docear's new digital functionalities? What is being performed? From these new practices and their perceived usefulness for those engaged in them, how can we reflect upon the changing relations between humans and academic text? This research tries to answer these questions. Its findings can be divided in two. First come the core empirical findings based on situated use; the details of practice both in and around Docear. Second are the user's perceptions of change.

## 6.1 Invisible Reading and the Lone User

In going from discussing my methods to analyzing my interview results I will begin with a finding that reflects on the methodological value of Docear. All of the users both discovered and worked with Docear independently. All had come across Docear via Internet searching and were alone in their use. All had recommended it to colleagues but none in their 4 to 18 months of experience with the software had shared common experiences with anyone in their personal or professional network.

At the same time, as will become clear in the course of this analysis, these users described Docear as having an observably profound impact on their use of texts considered central to their current research. All used it actively in research projects critical to their academic career. All but the MBA student mention how Docear is critical to helping them achieve their goals under time pressure. The use of Docear is not an ephemeral preparatory element to their work. Both The Assistant Professor in Communications and the Researcher

in Applied Nutrition describe how they do not have time to retrospectively integrate their library into Docear. They are rather using it intensively for the work at hand. I find this lone engagement to be a useful element for my research, especially when juxtaposed to these rather dramatic statements about the software's influence on their core work. Docear became a particularly revealing research object in conjunction with interviewees who, when combined with sustained engagement, both:

- Chose Docear independently and describe this as a personal experience of choice.
- Came to use Docear outside of a shared or collaborative setting related to a community of users with a set of externally imposed organizational standards or customs.

Below are excerpts of interview text that describe how the users discovered Docear. All of them found it by searching on the Internet. All of the users describe a preexisting interest in mind mapping that drove their impetus to find something new. All of them describe their discovery of Docear as serendipitous. They know why they chose it but not quite how.

Here are the responses to the question, "How did you find out about Docear?":

**Researcher in applied nutrition:** So I really read a lot about how Mind Mapping are really improving the way you organize your knowledge. And then I looked into software and then one day, I Googled "mind maps software" and "PhD". And then poof, all in a sudden, you get Docear.

-----

**MBA student:** I guess I was trying to increase my productivity, to move along some trouble, some issues, to make some progress. And probably I was looking at a website which offered several tips for increasing productivity.

---

**Researcher in applied biology:** I was looking as, thinking about getting something together for mind maps and I found a program called something Mind Maps and then it didn't really look that good. And I tried SmartDraw and I thought I was going to try and invent the wheel myself. I tried SmartDraw and I thought, "This is just not working. Somebody must have a better idea than this." And then I started Googling

and found Docear and I said, "This is exactly what I was looking for." It did all the work for me; it did all the stuff that I didn't want to do.

. . . chance, pure chance, I found Docear and started reading.

---

**Assistant Professor of Communication:** I have no idea. I was trying to remember. I know! Isn't that terrible?

...but [it] wasn't a direct recommendation from a colleague. It was somehow either Twitter or I saw a blog post somewhere or something like that. It wasn't from someone I knew.

. . . I guess I was doing the search for mind map. Okay, I think I was looking at what kind of mind maps software that was out there because I thought, "Okay, I'm visual and I need to try and do something. I want to try this as a way to organize my thoughts." And so I was looking at different mind mapping, free mind mapping software that was available to try out. And so I think when I was searching for that that I stumbled upon it.

---

**PhD Candidate in Intl. Strategy:** "I was stumbling around."

These descriptions of discovery describe Docear as a kind of independent object. The users' discover and adopt it with seemingly no outside influence other than "stumbling" upon it on the Internet. None had requested any custom support services prior to being interviewed. They had not participated in any sort of user group. They had worked solely off existing material made available on the web.

As described earlier, I am trying to overcome a challenge related to how one uses ethnography to make the consumption of text visible. About this Christine Hines says, "Users leave no marks on technologies, and interpreting the technology is often something they simply get on and do" (2001, p. 54). I will make the assumption that that "simply get on and do" in the context that Hines writes actually means, get on and do quietly and alone. A recognizable reality of the reading activities being considered are that they consist mostly of sustained solitary human interactions with text technologies. This is an important to

reality to identify. As we have seen both Hine (2000) and Borgman (2007) do already. If the users of Docear are lone users perhaps they will be more likely to be revealing those lone situated practices.

Because as lone users one would expect more varied use, it is particularly interesting that the use varied so little from user to user. This is despite the reality of a wide variation in skill and confidence that was revealed among the 5 interviewees. The Researcher in Applied Nutrition and the Researcher in Applied Biology were confident and definite in their use and both state that they would have found a way to create for themselves Docear-like functionalities if the software had not existed.

**Researcher in Applied Nutrition:** It's not only Docear; I mean if Docear would have not existed, I would have found another way probably much rustier to use maybe Xmind or freeplay, no other map... Mind map software to do the same thing but with Docear it's much smoother.

---

**Research in Applied Biology:** I was looking as, thinking about getting something together for mind maps and I found a program called something Mind Maps and then it didn't really look that good. And I tried SmartDraw and I thought I was going to try and invent the wheel myself.

Here Docear is described as addressing a need that they understood and had already anticipated. The PhD Candidate in International Business describes the opposite. Rather than being prepared to find something out on his own, he describes the need to learn to use the established functionalities of current text tools as a necessary part of his educational process.

**PhD Candidate in International Strategy:** So it's been 13 years since I've been in academic circles. Moving back into it now has been a painful transition. And I'm trying to make progress and learn techniques all at the same time, which is not helpful. I think... Had I... With hindsight, what I've have done is learn the tools really thoroughly to begin with and then structure all of my research based on the abilities of the tools. Now I'm in this transitional period of trying to do both at the same time.

These variations were not revealed as having a noticeable effect on how they adopted Docear or what they perceived as the usefulness and new value gained from using it. The interviews reveal relatively consistent use and perceived benefits among the 5 users. It is the combination of lone use and consistency of use among users that leads me to trust Docear as revealing something related to broader academic practices. And therefore, this use in turn impacts on that broader academic practice. Polanyi's independent scientific initiatives described in section 2 are "coordinated because each takes into account all the other initiatives operating within the same system." This study's findings show this to be a techno-social system where the medium of communication plays a critical role. Now the analysis will move on to descriptions of situated use directly related to digital reading practices that involve creating a personal library.

## **6.2 Distilling Actual Digital Libraries**

This section presents a basic element common to all the users' digital environments external to Docear. All of the users describe the creation of actual digital libraries. For the purposes of this study, digital text is made actual when a copy is saved and kept on a computer drive the user has control over. The ability to link an actual digital library of academic publications to Docear's mind mapping functionalities defines its core use for the five interviewees. This situated use will be shown later.

Four of the five interviewees used a reference management software to keep track of the bibliographic information related to various materials, regardless of format or whether they actually had the publication in their personal libraries. From my professional experience users tend not to link these bibliographic entries to actual documents, at least not in any consistent manner. The software is most often used to create a database containing primarily bibliographic (identifying) information about publications. It may link out to items

on the Internet and some digital things on your hardware, but it primarily acts as way to maintain control over bibliographic information. The name reference management suggests it manages references, which are virtual libraries as opposed to actual libraries of full text.

What is unique about Docear is how it changes the incentives to create and maintain actual digital libraries. It allows collected and stored digital PDF files to perform in new ways. This influences practice external to Docear. All of the 5 interviewed described, in relative detail, the process by which they download documents and include them in actual libraries on their personal system. They had a greater incentive to do this as an organized process because of how the documents would eventually link to Docear.

**The MBA student:** So what I do is I download everything first so I never do any online readings, too slow. . . . Second, I maintain my library in two ways, into separate sources. First, it's, let's say, just simple disk structure, the file system structure, where I have everything renamed according to a certain pattern. So I can always just use a Windows search. Second, if I feel that some article, let's say extremely relevant for my papers and I'm downloading and reading them for my paper, I will copy them into a separate folder which is a literature folder in my repository for thesis.

---

**The PhD candidate in International Strategy:** Right, 99% of my research is done through Google Scholar. So I'll go and search something in Google Scholar, find a document that I think, "Okay, that looks nice. I want to read it." And so I'll do a superficial read, abstracts and conclusion usually, and if I don't want to, I'll either import the citation information directly from Google Scholar into Endnote or I'll fill in the record manually. I then take the data record that Endnote[name of his bibliographic software] gives me and copy that. When I download the document from Google Scholar or whatever your database it is on, when I save the file, I paste it in as a filename the record reference from Endnote.

---

**Assistant Professor of Communication:** And I don't want to go back and go organize all those files right now that wouldn't be a very good use of my time. But as I use them to kind of build this base of organized material. Yeah, I'm trying to... I'm not sure I'm at right system yet. But the way I do it right now is the first part of the filename is the person's last name. The first author's last name. And then in the next, I'll go dash and then I'll put the year, so the year of publication. And then dash. These are kind of a... I might have a word or two that describes the contents of that

article. And then sometimes, I'll even put a dash at the end and an abbreviation for the publication it came from.

So what are the users of Docear doing in these situated descriptions of downloading and naming? They are systematically creating a digital collection of actual publications in discrete locations on their computer drives. One may argue that this doesn't sound like anything new happening. It is just the downloading of PDF files. Academics do it all the time you may say. "I have lots of them on my drives. Am I also distilling an actual library?" Not if you are not engaged in a systematic process of creating and maintaining a discrete, exclusively full-text, digital collection.

The norm outside of Docear is to store of academic literature in a more blended way. You might either store text digitally in PDFs, or just link out to a copy on the Internet. Sometimes you may print out a copy to read and other times you have purchased a printed copy. All of it is then managed through ultra flexible reference management software, which also includes many other items you do not have actual copies of, which are then virtual items. There may be PDF downloading going on in this set up, but this is not the distillation of an actual digital library. The personal library as managed in reference management software, is based on consolidated bibliographic information that is only in part represented by dispersed actual copies that are both print and electronic. The virtual is consolidated. The actual is not.

These relations are changed by the distillation process described in the user quotations above. The process extracts PDFs from existing reference management software and filters out all but full text PDF files from the Internet or elsewhere. An example of this is shown below in figure 5. It includes screen shots of both the author's file management software and reference management software.

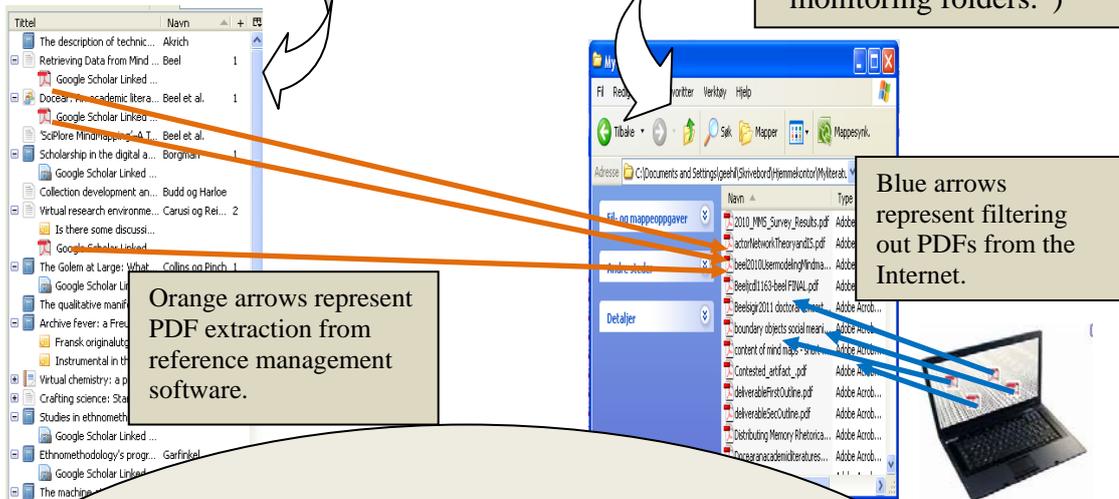
### Virtual library:

This is a title list from of a reference management software. It represents a virtual library that includes references to varied document types (note icons). It represents a potential library the author has not actually compiled.

### Actual library:

File manager title list from an actual library of digital publications in full text PDF format on the user's drives. (When Docear links to these libraries they are called "monitoring folders.")

 = PDF icon



## Distillation of actual digital text

This process was practiced by all interviewees as a basis for their use of Docear. Shown by the orange and blue arrows, it involves creating and maintaining a pocket of purely actual digital publications in the form of a folder of PDF files. This is done by either:

1. Extracting of full text PDF files from reference management software, represented by orange arrows.
2. Filtering out of anything other than PDF files from the Internet, represented by blue arrows

**Figure 5: Distillation of Actual Digital Text**

(Screen shots are from the author's reference management and file management software.)

Most reference management software can link out to PDF documents, as seen by the PDFs embedded in the virtual library in figure 5. One might ask, why don't users distill actual libraries here? Below I will use some ANTy anthropomorphizing to translate how a reference management software's obdurate, non-human qualities might argue against using it to maintaining a more rigid actual digital collection. It might say:

1. Here it is so easy to collect all kinds of things you might want to read; or listen to or watch for that matter. You may not have time to read this thing on your screen right now, but you don't want to forget it! Just click and include it in our database.

2. I don't have much of benefit to offer if you just have PDFs here.

And so, actual libraries don't happen. This is not showing itself do be due to the actual functionalities or the users plans. It is a normally silent negotiation in the hybrid network that makes the performance what it is.

It should be noted that this discussion of reference management software is based on few details of situated use coming from the interviews. It comes mostly from my experience as a librarian. But it is from descriptions of situated use that I mapped the distillation process shown by the orange and blue arrows in figure 5. And four of the five interviewed described bibliographic software as an element in the distillation process. In trying to better understand the new distillation process being described by the interviewees I was lead to reference management software as a dominant relation in the existing network.

## 6.3 Making Actual Digital Libraries “Your Own”

Another way to see how the actual digital gains ground due to Docear is to look at PDF use. PDF documents are often printed out. If the printed copy is the copy that is read and marked up it becomes the document they have actually used, even if they have a PDF saved on their hard drive. As the findings in this section suggest, due to the incentives of Docear, users read exclusively from the digital copy. PDF files become the text objects they keep and take care of and collect together; what I am calling distillation.

For four out of five users, Docear was described as the incentive to start using the highlighting and bookmarking features of a PDF viewing software. These four began doing this consistently only when inspired to use Docear. Only the Researcher in Applied Biology had been using them prior to using Docear. These types of changes to a PDF can be done only if one has an actual digital copy on a drive one controls and a PDF viewer software that enables the functionalities. These digital artifacts must reside on the computer drives that the user has control over, and one must have compatible hardware. There is no need, for an Internet connection to do these things. One could conceivably transfer both PDF files and PDF software without using the Internet. That said, the fact of the Internet with its searching, connecting, and downloading functionalities has set a precedent for creating and working with digital libraries in Docear. The Internet has made this whole system viable. The ability to work offline with what one has collected locally does not change this.

Here are descriptions of the situated use of PDF viewers. They describe new reading practices that are external to but attributed the use of Docear. However, these four users attributed these new PDF practices to having started using Docear. (Note regarding

subsequent quotations: in case it is not clear out of context, all of the below quotations describe working with and reading in an exclusively digital arena.)

**The Assistant Professor in Communications comments:** When I started using Docear and for the first time, I hadn't been using bookmarks or highlighting in Acrobat before. So when I saw that Docear could incorporate that, I said I want to figure out how to do that in Acrobat, I thought this was helpful.

-----

**PhD candidate in International Business:** Foxit is the PDF [viewer]... Don't call it an editor because I've got the reader version. But it's great as an editor too. It gives me enough in terms of reading academic papers, highlighting, bookmarking, commenting, and other bits and pieces. It has been fabulous.

-----

**MBA student:** if I feel that some article, let's say extremely relevant for my papers and I'm downloading and reading them for my paper, I will copy them into a separate folder which is a literature folder in my repository for thesis. Then I would, let's say, bookmark them and arrange sources and arrange phrases and then I will import them into my Docear.

-----

**The Researcher in Applied Nutrition.** What I do... Okay, in sequence, I know that I have to read that certain amount of papers in my phase of literature review. So, I just scan one by one all the papers. I read them very thoroughly. And every time I find a good idea, a good input for one of the components of the papers and writing or the chapter I'm writing, I simply highlight it as a bookmark and that's it.

The distillation of actual digital text is reinforced by the use of PDF view software. Digital files stored on their drives are theirs to keep and personalize. We begin to see situated uses of changing infrastructure in the relations between a group of individual scholars, file types, and software compatibilities. The four quotes directly above reveal new uses of existing PDF software beyond simply viewing and storage. These new practices were described as side effect of users establishing themselves on Docear. It is not until the last two quotations above that we begin to hear descriptions of the importing of files into Docear. Both the distillation of actual libraries and this new uses of PDF viewers show

Docear practice extending outside of its own boundaries. It is not until the linking to mind maps begins that situated uses of Docear mind maps is revealed.

## 6.4 Visually Structuring Actual Digital Libraries

Now I will proceed with descriptions of the situated use of Docear mind maps. It takes place within localized “actual digital” environments that the interviewees have configured for themselves, as described in the preceding sections. All of the users describe importing their PDF files into Docear. Not to get too technical, but this is perhaps more accurately described as hyper-linking to what Docear calls “monitoring files,” which contain their actual digital libraries. This process is about mind mapping not only ones’ own ideas but also the ideas of others as they exist in the text of separately published PDF files. To begin describing how this is performed and related to both the distillation and annotation processes, I will start with the repetition and then continuation of a previously-used quotation.

**Researcher in Applied Nutrition:** If I want books or papers, I order them, I download them online and I keep them digital; I don't print them anymore. And every time I find a good idea, a good input for one of the components of the papers and writing or the chapter I'm writing, I simply highlight it as a bookmark and that's it. And then... Later on, I opened Docear and I import them. So one by one, I transfer every single idea into what I think they should be their final location in my paper.

The “good ideas” he refers to here are other people’s ideas found in texts he keeps in PDF format on his personal computer drives. The final location in his paper at this point is not a word processing document but a mind map of the ideas for his paper, his and others’. As mentioned earlier, new ways of reading existing academic texts are not easily to separate from new ways of writing original academic texts. However, this research makes an attempt and the above quote is a clear description of the chain of academic reading practice surrounding Docear.

1. Importing relevant PDF publications into one's actual digital library, this is part of a continuous distillation process related to the situated use of Docear. ( The search/evaluation process is not dealt with in this research)
2. Personalizing the text in the digital library with the bookmarking and highlighting features of a PDF reader software.
3. Dynamic linking (often called importing by the users) of text excerpts or entire documents from one's actual digital library into a Docear mind map.
4. Using Docear's features of linking mind maps to PDFs is the basis for it being an integrated tool for note taking and organizing the content of PDF-based text content. Notes, entire publications, and fragments of publications can be mapped in relation to one another and combined with notes.

Point four describes organizing and structuring text in a library. These are the capabilities that the users mention over and over in different ways as a the uniquely useful reality of using Docear. This system is described as an aided not just in structuring thoughts but in structuring thoughts around a discrete body of literature. Books on a shelf can be reorganized and mix up, Docear enables such structuring but not just of entire texts but also excerpts from those texts. In addition one can replicate different combinations, like having many bookshelves of the same books arranged in different manners.

I want to interject here that many types of ideas and many types of documents can be linked to mind maps. One can also embed Internet links and link out to documents available on the web, but none of the interviewees mentioned this type of use. Situated use related to actual digital PDF documents dominated my conversations with all of the users. And when I asked for inventories of their Docear mind maps it was clear that this dominated their situated use; apart from some who used it for administrative tasks unrelated to managing or

reading external text. Similar to what was described earlier with the use of reference management software, it is not about plans and existing functions but a situated negotiation between the software and the user. Unlike reference management software, with its strengths in relation to creating virtual libraries, Docear had a stronger argument for keeping things actual.

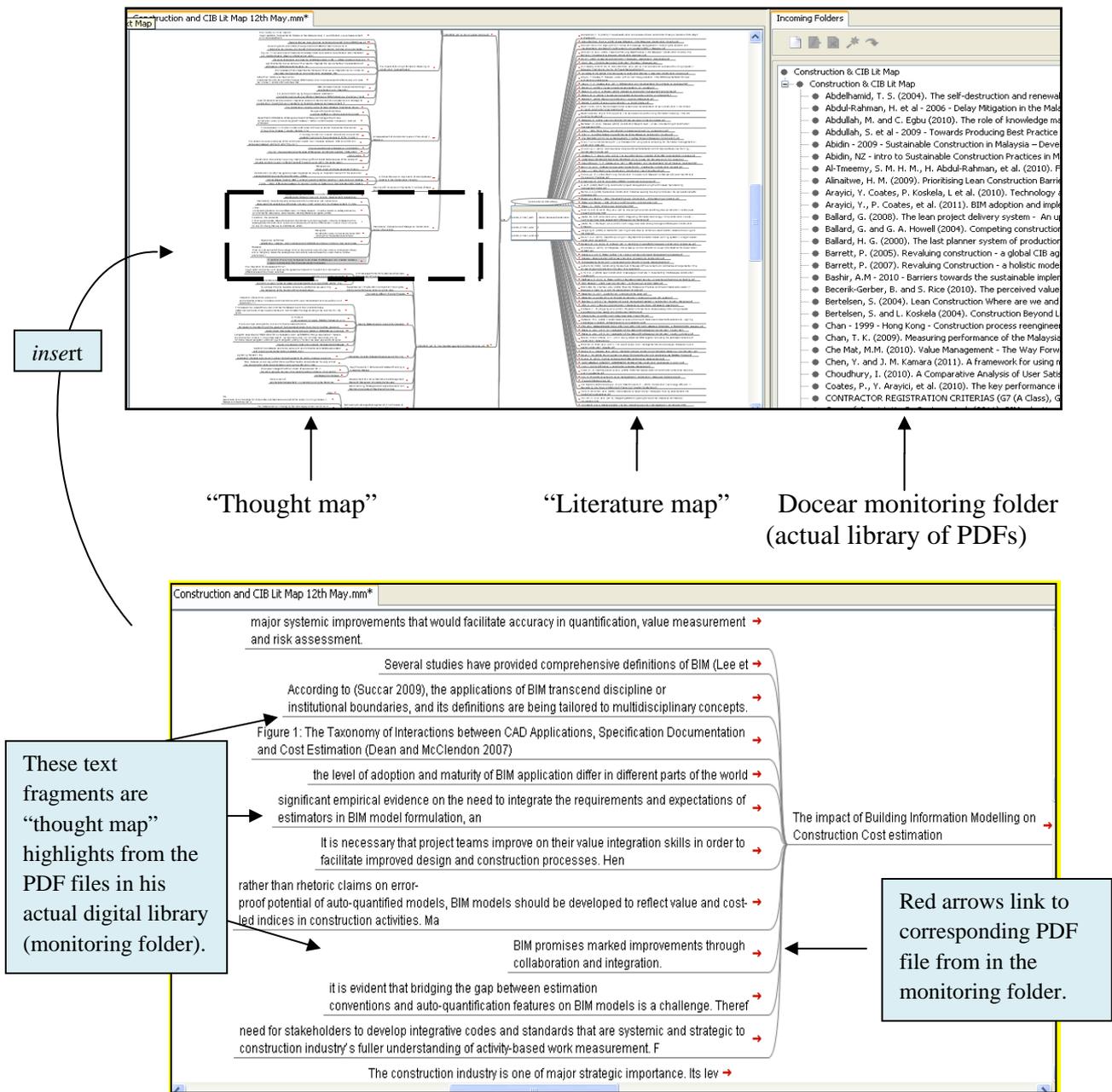
This quote from a PHD candidate in International Business reveals this argument.

**PhD Candidate in International Business:** So I bookmark... As I read through the paper, I bookmark it as necessary. Then I upload it onto Docear so I can see all of my literature there with all the comments in there, and I can copy and paste them from one to... From my literature map into my thought map as such.

In addition to this brief description, opposite is a Docear screen shot of one of this user's mind maps, see figure 6. In the middle of the upper thumbnail image is a mind mapped list or "literature map" of links to PDF files he has in a folder on his computer. The contents of the folder itself are shown on the far right. This is called a "monitoring file" in Docear, which is an external folder containing PDF files that generates an initial mind mapped list of links to one's digital library.

Links to the contents of monitoring files appear automatically in a Docear mind map when new files are added or existing files are highlighted or bookmarked. The mind mapped list can be set up to update in alphabetical order by filename. (You cannot see this in figure 6 since capturing the map's structure required reducing the image to illegibility.) I infer that this alphabetization feature explains in part the attention to file naming conventions the interviewees described earlier. In addition to a hyperlink to the PDF where it resides, each

article represented on a mind map is a node that can open, or “unfold” to reveal bibliographic data, bookmarks, and highlights. This is shown more clearly in the insert zoomed in on in figure 6.



**Figure 6: Working with an Actual Digital Library in Docear**  
 (note: Screen shots from the PhD Candidate in International Business. “Thought map” and “library map” are his words.)

The screen prints in figure 6 as described in the preceding quotation exemplify a structuring that all 5 users describe with some degree of variation. He created a mind map from a Docear monitoring file; which serves exclusively as an overview of the PDF documents residing on his computer drives. This is what above user calls a “literature map.” His “thought map” is where he selects and arranges nodes from his “literature map” and combines them with notes and organizes them around a structure related to something specific he is writing. As he copy and pastes sections of mind map within and between “literature” and “thought” maps the connection to the copy in his monitoring file is preserved. Updates affect all the copies/links in the mind maps. Here is the MBA student’s description of this process using the same features.

**The MBA Student:** Docear, that's the only thing you need. Because it helps you to give a... You just drop a set of key words, then you rearrange the sources and extractions of bookmarks and text from PDFs according to this, let's say pattern. Then you just copy and paste the whole tree, do some editing stuff and here you go; these are the concepts you need to handle. And more important, that's the way you track the references. . . . So, the first application, I use it directly as a mind mapping tool. That implies structure and ideas into hierarchy, all those fancy folding functions and drag and drop features. That's pretty straightforward, just mind mapping tool. Second is batch extractions of bookmarks from PDF in an organized way. Third is structuring extracted text, texts into personally defined categories.

Susan Leigh Star says that as regards technological infrastructures, “developments move either from independent or dependent variables, to processes and relations braided in with thought and work” (1999, 380). Here we begin to observe Docear moving from being an independent variable whose existence and qualities have little or no bearing on its surroundings to a central element in the working lives of these five scholars. Their situated use is not random. We see repeated activities forming practice. That these practices are common to all 5 lone users gives credence to the practices themselves being a form of maintaining some accountability to external, “relevant professional and organizational constituencies” (Suchman Trigg Blomberg, 166).

The challenge given to “indexicality” in this research is to discover things discussable in Docear’s new reality that reflect new conventions related to reading academic text. Similar to Suchman Trigg and Blomberg’s prototype, Docear is a new tool out for a test and has shown the potential to “illuminate the coherence of technical artifacts as contingent achievements” (2002). An understanding of its agency within the infrastructure of academic literature is derived neither from user needs (plans) or the functionalities of Docear. Needs are not originating in the user nor the software, but rather in the performance. This becomes especially clear in findings described in the next section where users describe their experiences of space collapsing between texts in their digital libraries.

## **6.5 Space as Connection in Actual Digital Libraries**

The next two sections relate to where the interviewees offered descriptions of the ways in which the new use and detailed academic reading practices described so far represent change. I will begin to move away from analysis based solely on descriptions of situated circumstances of use. I begin to present more of the interviewee’s own reflections on what has been happening. However, the base of situated use built up around these discussions is in contrast to Garfinkel’s hopeless work-a-day conversations since the speakers here are in the midst of a technical project of sorts. There is an indexicality built up between the interviewee and myself as we discuss broader changes related to their use of digital academic texts in Docear. I asked questions related to their reading preferences for electronic text vs. printed text. All five interviewed preferred to read academic text in electronic format. The applied nutrition researcher made these definitive statements related to perceived changes since using Docear

**The Researcher in Applied Nutrition:** Before, I thought it was inconceivable to read something not in print and then retaining things. While now it's the opposite. If it's a... The capacity of taking notes of every single... concept yet it comes up during the reading through Docear and then making the bookmark and so forth. Well, it's ten times richer... when you write something, you just scribble something. Then you forgot what you tried to scribble, to know exactly what you meant. Well, using both the bookmark and the location on the map of Docear, then it's perfectly clear what you meant and in which context you want to use that piece of information in the future.

. . . Actually, since I discovered Docear, I completely cut down any printing... Okay, absolutely no loan of papers, absolutely no loan of books. If I want books or papers, I order them, I download them online and I keep them digital; I don't print them anymore. But one of the secrets was Docear, because before I had to print them and read them and scribble on them, but not anymore.

He describes trying to understand what he has written “scribbled” on printed versions of academic text. After the fact he has trouble integrating it into his working notes and outlines. It is not just the problem of hasty handwriting but also location, where and how it relates to other texts. In Docear’s new “infrastructure of relations” he identifies a conundrum with printed text. The physical distance between them, no matter how short, is a now a barrier to integrating their ideas and finally writing text based on those ideas. These distances, even within one’s own office, becomes too great when compared with PDFs in Docear where the article and its critical passages are instantly connected to each other and his notes. He had built up a library of approx 1000 academic publications in PDF format and of all interviewees had been using Docear the longest, 18 months. He was one who stopped reading from printouts after starting with Docear.

The Associate Professor of Communications also describes how her relationship to reading digital documents had changed since using Docear. She is one of the three users, including the Researcher in Applied Nutrition, mentioned directly above, who give Docear credit for a new preference for reading academic literature electronically instead of in print.

They refer specifically to academic publications in their own “actual digital libraries” linked to Docear functionalities.

**Question:** So when you're reading your academic or scientific literature, do you prefer to read it in print or [electronic format]?

**Assistant Professor of Communications:** That's in flux, right now as soon as like maybe six months ago, I would have said paper, definitely. But especially with my discovery of Docear and using that to help me with the literature review, I've been doing more [electronic]. So I think I've been making the transition from preferring paper to preferring [electronic format]. I think I retain a little more when I read it by paper. But having the organization and the access capabilities of working with electronic documents, I think I'm probably going to do more of that now.

[When discussing her use later on she describes this organization a bit more:]

. . . And then it's really nice to always feel to get to that same place online, instead of this piece of paper. For example, I was looking at [a] citation today, I was like, "Oh, I want to go look at that original source again and see something", and it's something I don't have an electronic [version] for. Those things I do now, and started to think about where that might be, and how to get to it. Well really, I just want it at my fingertips right now because I'm getting spoiled.

[She goes on as I try to draw out more description of this experience of Docear.]

. . . Speed, synthesis and then having multiple locations where I work. It's having one place that I can go, instead of... If I have a physical hard copy, it's going to be somewhere. It could be maybe at my house, or maybe in my office, or maybe in a bookstand somewhere. Where is it that's electronic, it's just right there.

She describes the inverse of what The Researcher in Applied Nutrition described in the previous quotation. She is starting with the notes and trying to get back to what is referenced there. She is “spoiled”. She wants it at her fingertips. Space in her Docear mind maps is collapsed by connection (Hine 2000, 106). She is coming to expect zero distance between the place where she is writing down her thoughts and her library.

### **Connecting expanded and contracted text**

It is not just links from one area of text to another. Change in Docear is also described as being a result of instant connections between expanded and contracted visual landscapes of text. The MBA student describes situated use of expand and contract features

in Docear's digital mind maps, this feature is shown previously in figure 2. He begins by explaining challenges related to the sheer volume of text being considered.

**MBA student:** By seeing what has been done in every single field, you're able to determine where exactly you should have to place your efforts. . . . I will tap on this, and then I will write more about this, and that will become my conceptual model. And here, Docear helps you first to see. Let's say on the left you have all that's been covered, on the right what has not been covered. It's a visual representation. I don't know if it's... Sometimes I'm feeling like... Well again I'm done with stages.

[He laughs here as he explains that he has finished a description he has prepared for our interview. I assume what came after this point was not prepared.]

Sometimes I just feel it would be of great help if I could arrange every single piece of needed information on one screen.

**Interviewer:** So you need a bigger screen?

**MBA Student:** It's never big enough, you know? So, sometimes you really need to have an ability to fold the text, like it happens in Docear.

**Interviewer:** To fold, yeah ... the way things can collapse and open in a mind map?

**MBA Student:** That's true. So in that way, you can do let's say, some statements, assign them one label that's self explanatory, and then fold the rest of things.

[With a bit prompting to explain his decision not to change screens he says:]

. . . but the bigger it becomes the more you feel that there is nothing compared with your brain with the power of your mind. When it comes, let's say, to seeing things in general. The screen doesn't help in that. . . . [Docear] helps your brain to see the picture in more general, so after all you don't need a bigger screen.

We are seeing in this chapter that for the interviewees their actual digital library is “right there” for them in arena created by Docear. Users are creating instant links between entire texts, text excerpts, and their own notes. Mapping allows for new visual arrangements and juxtapositioning. Complex landscapes are created where extensive branches of text and links can be expanded and collapsed within a single-screen working space. Space is entirely defined by connection; not any distance apart from the screen itself (Hine 2000, p. 106).

### **e-Reading's Tipping Point**

All of five Docear users preferred reading academic text in electronic format over print format and three of gave Docear credit for passing this tipping point. They are “spoiled.” Things they accepted before in print are “inconceivable.” This reveals that the digital is evolving into an arena for maintaining actual academic libraries for personal use—an arena where the full-text of publications are kept and read. This stands in sharp contrast to what are the common practices described earlier where personal actual collections are spread over various mediums while the personal digital arena is what holds it together bibliographically or virtually. As users describe what is happening within these new relations between space and text they use words like: “tracking”, “retaining”, “clarifying”, “recalling connections”, “seeing connections”, “seeing relations.” Something has changed that is beyond mere practical advantages of storage and retrieval.

## 6.6 Tangibility of Actual Digital Libraries

In trying to interpret Docear users’ descriptions I collected some of their words that reflect work with the actual: “tangible”, “concrete”, “workable”, “capabilities of working” “scale” “purely controlling”, “synthesis”, “mind dump”, “structure.” Docear’s own use of the word, “structure”, seems to point to what is happening with this change. The ethereal virtual digital is made into a concrete actual digital. Here I present more comments reflecting this perception.

**PhD Candidate in International Business:** I think what I find that helps me is, once you've done the mind dump and put it onto a map, you can let it stand back and look at it objectively. You're no longer lost in this game of trying to keep memories alive.

-----

**Interviewer:** Do you feel like your literature is organized in a new way since using Docear?

**The Researcher in Applied Biology:** It’s more tangible for me now, I believe.

**Interviewer:** Tangible in what way? Can you explain what that...

**The Researcher in Applied Biology:** I can see it more clearly on how things pertain to each other. Because of the subject matter that I'm working with, it has the particular receptor that I'm working on has a lot of effects. And it is difficult to try and keep track of all of them and interaction of all of them and with a mind map it's making it actually to sit back and have a look and, yes I can see now that's how each one is relating to the other. [I worked to have him draw this out a bit further] . . . The actual receptor itself has over 25 different areas of where it had its effects. So it's a fairly extensive, so just keeping track of all of them and new ones coming on board all the time it was getting a little difficult to see without a mind map; that's why I wanted to use a mind map.

By tangible, he does not say that he realizes more than ever that there are bits of digital material on his computer that make up the article he is reading. He rather says that he has more clarity regarding the ideas presented in digital text and particularly how different digital articles' ideas are easier to relate to one other. But one could argue here that what is happening is a physical mastery of the digital medium. Here is The Assistant Professor of Communications.

**The Assistant Professor of Communication:** Using Docear helps me to do a better job of synthesizing and bringing ideas, seeing connections, then bringing ideas together and seeing links that I might not have seen before. . . . So before, maybe I was summarizing more than synthesizing. I'm doing a better job at synthesizing because of the mind map function of it. . . . So probably elevated my scholarship, I would say.

She uses the word synthesize and we can easily associate that word with purely cognitive processes, but we know in this context that links, connections and bringing ideas together is also happening physically in the form of manipulations of actual digital text.

When something is distilled its unique qualities become distinct and observable out from a blend of materials. Its properties are better understood and used. At first, all these comments about the tangible and the concrete don't seem to be about a literal sense of the

actual digital bits of material on their machines, but perhaps in some way they are about that. Users suddenly had complete digital artifacts in complete form; undisturbed by the handicap of mere speculation that the virtual always introduces. The virtual in the ontology of this research can only speculate about what the physical presence of a particular text might mean for understanding. “What would sections of that article do pasted in here do for my understanding of the topic?” One does not know if one does not have it there. The digital publications being structured in Docear are kept, physically. Physically (on a digital level) things are being altered, moved, manipulated, partially replicated, and structurally related to each other in a freer way than perhaps ever before. New physical standards for performance are set for text. Relations between and among texts are enhanced. New special connections and perspectives are gained over discrete bodies of literature.



# 7 Conclusion

Ideas expressed by the file include the notion that human expression comes in severable chunks that can be organized as leaves on an abstract tree—and that the chunks have versions and need to be matched to compatible applications. (Jaron Lanier 2010, p12)

raw experience . . . the elements in it are verbally fixed and coupled together so that we might know what is in the wind for us and get ready to react in time. (William James 1912, p. 96)

Like Marshall McLuhan 50 years ago, in the above quotation Jaron Lanier also requests that we see the “medium as the message” or at least a message. Also hailed as a media guru of sorts he too uses the broad brush of the macro. He speaks here of more than just text, but it is still relevant to the findings of this research. It’s about the physical nature of the digital and its “struggle” to be present (actual) in order to perform. It also speaks to the importance of how texts are relationally structured. The William James quote, originally published in 1905 supports Lanier’s statement that file systems have roots in an earlier idea, “raw experience . . . the elements in it are verbally fixed and coupled together so that we might know . . .” Raw experience according to James is not useful simply fixed through text. It must also be coupled, partnered, and arranged in relation to other texts. That is part of its performance.

The discoveries of this thesis reveal a reconstruction of this performance in personal shifts in how virtual and actual libraries manifest themselves. We can now return from the language of this study’s ontology to daily-life descriptions. We see the five Docear users shifting their academic reading practices entirely into the digital arena. They all read and work with academic text in a digital format instead of just searching and managing it on the

computer and then reading in some form of printed document. This research reveals how this changes how they work with text and what sorts of new reading practices are evolving as a result. It revealed readers using Docear to facilitate working with digital texts in ways that make visible what James and Lanier describe—“coupled together”, “organized like leaves on an abstract tree.”

The implications of academic reading moving entirely into a digital arena are far reaching. Jaron Lanier certainly does not have utopian stars in his eyes when he uses the word “chunks” of information. He is involved with groundbreaking work within digital mediums at the same time concerned about how its nature has entirely new sorts of lock-in or deterministic tendencies. “Software is worse [as a tyrant] than railroads, because it must always adhere with absolute perfection to a boundlessly particular, arbitrary, tangled, intractable messiness” (2010, p 258). This STS research in no way closes the door to these concerns. Neither does it promote a particular product or practice. Neither does it negate the author’s personal sadness and anxiety about what she perceives as the gradual demise of the printed book.

That said, this study’s results are limited to micro considerations of the performance of academic text for a handful of readers. Here digital text is seen having found a degree of resolution for its “chunks with versions needing to be matched with compatible applications.” Consider today’s grainy, work-a-day world of actual academic text with printed works, computer printouts, proliferating e book formats, PDF files plus multiple reading and storage devices. Out from this, the results of this study show the PDF/Docear pairing of standardized, digital publications and mind maps to stand as a particularly potent reality within the digital evolution of academic reading practices. Through the digital collapsing of space readers are drawn deeper into their own library of actual digital text.

This in turn can make virtual text out of a printed book on a shelf in the same room. Its passages on paper are just not right there to be organized together with other text's passages "like leaves on an abstract tree." In this way that we see libraries and reading itself reconstructed in a way that makes the digital text more tangible than text on paper. This is not a fixed conclusion but a revealing perspective built up and interpreted for anyone doing work related to academic text. It can be used, as William James put it, "so that we may know what is in the wind for us and get ready to react in time"; in this case to things in the practice of reading that reside just outside the frame of current inquiry.



# List of References

- Akrich, M. (1992). *The Description of Technical Objects*. In Bijker & Law (eds) *Shaping Technology/Building Society*. Cambridge, Mass.: MIT Press.
- Beel, J., & Langer, S. (2011). An Exploratory Analysis of Mind maps. *Proceedings of the 11th ACM symposium on Document engineering*. New York: ACM.  
doi:10.1145/2034691.2034709
- Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. London: MIT Press.
- Bowker, G. (1994). Information Mythology: The World of/as Information. *Information Acumen: the Understanding and Use of Knowledge in Modern Business*, 231–247.
- Budd, J. M., & Harloe, B. M. (1997). Collection Development and Scholarly Communication in the 21st century: From Collection Management to Content Management. *Collection Management for the 21st century: A Handbook for Librarians*, 3–25.
- Collins, H & Pinch, T. (1998). *The Golem at Large: What You Should Know About Technology*. Cambridge: Canto.
- Collins, H., & Yearley, S. (1992). Epistemological Chicken. In Pickering (ed.) *Science as Practice and Culture*, 301–326. Chicago: Univ. of Chicago Press.
- Deleuze, G., [1968]. (2004). [English translation by Paul Patton ]. *Difference and Repetition*. London: Continuum.
- Denzin, N. K. (2010). *The Qualitative Manifesto: A Call to Arms*. Walnut Creek: Left Coast Press.
- Garfinkel, H. (1994). *Studies in Ethnomethodology*. Cambridge: Polity Press.
- . (2002). *Ethnomethodology's Program: Working out Durkeim's Aphorism*. Oxford: Rowman & Littlefield.
- Guédon, J. C. (1999). The Digital Library: an Oxymoron? *Bulletin of the Medical Library Association*, 87(1), 9-19.
- Hine, C. (2000). *Virtual Ethnography*. London: Sage.
- James, W. (1912). *Essays in Radical Empiricism*. New York: Longmans, Green & Co.
- Lanier, J. (2011). *You Are Not a Gadget: A Manifesto*. New York: Vintage.
- Latour, B. (1999). *Pandora's Hope-Essays on the Reality of Social Science*. London: Harvard University Press.
- . (2005). *Reassembling the Social: An Introduction to Actor-network-theory*. Oxford University Press, USA.
- Latour, B., & Woolgar, S. [1979]. (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton; N.J.: Princeton Univ. Pr.
- Latour, B., & Yaneva, A. (2008). “Give Me a Gun and I Will Make All Buildings Move.: An ANT's View of Architecture.” *Explorations in architecture: Teaching, design, research*, 80–89.
- Lundgren, L. J. (2000). “The Utilization of Research. I Lundgren”, Lars J. (ed), *Knowing and Doing. On Knowledge and Action in Environmental Protection*. Stockholm: Swedish Environmental Protection Agency.
- Marcus, G. E. (1995). “Ethnography in/of the world system: the emergence of multi-sited ethnography.” *Annual Review of Anthropology*, 24, 95–117.
- McLuhan, M. (2003). *Understanding Media: the Extensions of Man*. Corte Madera, Calif.: Gingko Press.
- Nicolini, D., Gherardi, S., & Yanow, D., ed.(2003). *Knowing in Organizations: A Practice-Based Approach*. London: ME Sharpe Inc.

- Polanyi, M. (1958) *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: Univ. of Chicago Press.
- . [1962] (2000). The Republic of Science: Its political and Economic Theory. *Minerva*, 38(1), 1–21.
- Star, S.L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, 43(3), 377–391.
- Star, S.L., & Ruhleder, K. (1994). Steps Towards an Ecology of Infrastructure: Complex Problems in Design and Access for Large-Scale Collaborative Systems. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, CSCW '94* (s. 253–264). Chapel Hill, North Carolina, United States: ACM. doi:10.1145/192844.193021
- Suchman, L.A., Trigg, R., & Blomberg, J. (2002). Working Artefacts: Ethnomethods of the Prototype. *The British Journal of Sociology*, 53(2), 163–179.
- Suchman, L. A. (1987). *Plans and Situated Actions: the Problem of Human-Machine Communication*. Cambridge: Cambridge Univ. Press
- Wheeldon, J. P., & Faubert, J. (2009). Framing Experience: Concept Maps, Mind Maps, and Data Collection in Qualitative Research. *International Journal of Qualitative Methods*, 8(3), 52–67.
- Woolgar, S. (1991). Configuring the User: the Case of Usability Trials. In Law, J (ed.) *A Sociology of Monsters Essays on Power Technology and Domination*. London: Routledge.
- Wynne, B. (2003). Seasick on the Third Wave? Subverting the Hegemony of Propositionalism: Response to Collins & Evans (2002). *Social Studies of Science*, 33(3), 401–417.

## ***Appendices***

# Appendix 1: Template for Introductory Letter and Terms of Agreement.

Dear \_\_\_\_\_,

Thanks so much for agreeing to be interviewed about your use of Docear! I am looking forward to meeting you via Skype or telephone. Below are 4 practical questions regarding technology and scheduling. A 5<sup>th</sup> question is to accept the terms regarding interview data. (The terms themselves follow question 5). It is very important to me that you understand and are comfortable with how I will be using your interview data.

Please answer these questions at your closest possible convenience. Just write your answers in the spaces provided below and return. Of course feel free to include other questions or comments. For reference, here is the link to SciPlore's blog entry about the interviews: <http://sciplore.org/blog/2011/04/20/be-a-guinea-pig-take-part-in-a-user-study-and-help-improving-sciplore-mindmapping/>

Best Regards and thanks again!

Hilah Geer

Oslo, Norway

ESST Masters Student, University of Oslo

## **Docear Interviews: Times & Terms (Answer in the spaces provided and reply via email. Thanks!)**

1. What are three possible 30 minute time slots that are good for you? (I will be choosing only one. Ideally, these will be weekdays between May 9<sup>th</sup> and May 20<sup>th</sup>. If necessary I can do weekends and conduct interviews through the end of June.) Your answer:
2. What town or city will you be in during the interview times mentioned above? (This will enable me to take time zones into consideration.) Your answer:
3. Do you have Skype? What is your Skype name? Your answer:
4. If you do not have Skype, what telephone number can I reach you at for the interview? (Keep in mind that a mobile phone might incur costs to you even if I initiate the call. Choose a land line if you have access to one or check with your mobile provider. I will be calling from Norway. ) Your answer:
5. Do you both agree to and feel comfortable with the terms regarding interview data(see terms below)?  
Your answer:  
Yes\_\_\_\_\_ No\_\_\_\_\_ (Please explain so that I might set up interview data terms that you are comfortable with.)

## **Terms related to Docear interview data gathered by Hilah Geer in May and June of 2011**

- i. Unless otherwise specified by the interviewee, all interviews will be recorded. These recordings will be erased after Hilah Geer's resulting thesis has been evaluated.
- ii. These recordings and subsequent transcription data will be housed exclusively on a secure university drive or a password protected personal computer.
- iii. All information gathered from the interviews will be kept entirely confidential. Only people in an advisory role in an academic setting will have the opportunity to see any identifying information. Otherwise, no identifying information will be shared. At no point will identifying information be shared with Docear.
- iv. Unless otherwise specified by the interviewee, all processed anonymous data will be shared with Docear in the hopes that it will be of use in their further development of the software.

# Appendix 2: Interview Script

Interview script page 1/1

## Docear User Interview Script(undistributed)

I will start with a brief explanation of my plan for this interview. Feel free to ask me questions at any point and let me know if there are questions you would rather not answer. All of these questions will be made anonymous when I process the data, proper names such as cities and institutions will not be incorporated in the data.

I will be asking a range of questions. I will begin with specific questions that that help me understand the context in which you work; demographics, area of study, technological context. I will continue with a few open questions with which I am hoping to hear a bit of the story of how and why you use Docear in your work. Then I will be concluding with questions related to SciPlore's specific functionalities. It is useful for my work to learn in general how you work with literature and text so it is fine at any point if you find yourself discussing activities outside of Docear.

If this sounds okay to you and you do not have any questions at the moment, I will proceed. Is it important for you that I stop our interview in 30 minutes?

I. Context questions: We do not need a lot of detail.

Work / education:

1. In what field do you work? If you are involved in research or teaching can you describe the specific area?
2. What institution(s) do you work? Can you briefly describe your position?
3. In what town or city do you live or work?
4. Can you describe your formal education? Do you have any higher education degrees and or are you currently enrolled in a degree program?
5. Is it okay if I ask your age?

Technology:

1. Can you describe/name the different hardware you use in your work? Laptop, stationary machine, mobile phone, reading device?
2. Can you describe what sorts IT/Library services you have access to?
  - a. VPN, networked drive, cloud computing.
  - b. Library services, type of access and remote access, database access
  - c. Use of library services to order/access literature not available to you online.
  - d. IT support for software and hardware
3. Can you take a minute to describe the physical breadth of your personal research library? I will define this as the collections of copies that you personally manage for your research or academic work. This could include printed materials and electronic materials in various locations. Locations would include mobile phone, reading devices, bookshelves at home and at work, laptop, network drives.
4. When reading academic/scientific literature do you prefer to have it in print or online? What do you do most often? How many different devices do you use for this reading?

## II. Open Questions

1. How long have you been using Docear?
2. Can you tell me about how you have used Docear recently, say in the last month? [If they have not used it in the last month I will ask to describe the most recent period of active use.]
3. How often do you use elements of Docear in the course of a working day?
4. Have you seen the Docear video? Did this video inform your use and how does your use of Docear compare to the example shown in the video?
  - i. ([http://sciplore.org/software/sciplore\\_mindmapping/](http://sciplore.org/software/sciplore_mindmapping/)) 7min.
5. How did you find out about Docear and what were your reasons for beginning to use it?
6. How was it to get started with Docear?
  - a. How much time did it take before you were able to start using some of the elements of Docear. What was frustrating what was fun or interesting? Did you get any help from Docear or other colleagues using Sciplore?
7. Do you know anyone else using Docear? Have you shared experiences? Have you had any experience with using Docear collaboratively?

## III Functionality Questions:

Now I am going to go through SciPlore's main functionalities but ask you general questions about these processes. I am interested in how you go about this processes and the degree to which Docear plays a role.

1. About how many Docear mind maps do you have?
2. Which one have you used most in the past month?
  - a. Can you describe it and the last few times you used it?
3. Do you link your Docear mind maps to bibliographic references.
  - a. Have you created a large relatively inclusive mind map of your electronic library?
  - b. Does this also include references to literature that you do not have stored in fulltext on your hard drive?
  - c. When you reviewing your literature or looking for a document on your computer, do you use a Docear mind map?
4. How much of your personal library would you say is incorporated into SciPlore's mindmapping system? How much of it has been highlighted?
5. Have you linked Docear to a word processing software? Did you use it actively for the last piece of academic writing or something you are currently writing?
6. Do you share copies of your mindmaps with others? Can you explain how and why?
7. Do you use a bibliographic software like Zotero, Mendeley, or Bibliofil. Is it linked to Docear?
8. Is your literature organized in new way since establishing a Docear database? Describe.
9. Imagine your personal research library was entirely electronic and you had no access to print resources. Would it be a difficult for you with your current set up and equipment to make optimal use of this collection? Explain if you can.

Thanks so much for your willingness to give of your time and experiences. Jöran and I will be sure to send you information about any articles that might result from this research. If in retrospect you have other things you would like to share feel free to send me additional comments. Copies of mind maps are welcome. These of course can be very informal and perfect English is not important. Do you know of anyone else who might be interested in being interviewed?

## Appendix 3: Scribie's Transcription Process

Description of Scribie's Transcription Process

### Transcription Process

We follow a completely manual transcription process and manage it end-to-end. We process each file in four stages.

- **Distribution:** The file is split into smaller parts and distributed to our transcribers.
- **Transcription:** Each part is transcribed into what is called the Raw Transcript. The portions which are inaudible and/or incomprehensible are marked with blanks at this stage.
- **Review:** The transcripts are reviewed, mistakes are corrected and blanks are filled. Timestamps and speaker initials are also added at this stage. The remaining blanks are timestamped.
- **Proofreading:** All parts are collated into the final transcript, inconsistencies between the parts are corrected and the final transcript is reviewed once more thoroughly.

Our system manages all the above stages. It also takes care of recruitment of transcribers, monitoring their quality of work and making payments to them. You do not have to get involved in evaluating, selecting and monitoring as is generally required on other freelancing sites. Of course, you can check the transcript as it is being prepared anytime from your account.

The transcribers are vetted thoroughly before they can start working with us. They have to appear for a test initially and all their work is reviewed and rated and they system monitors their performance continuously. We provide them with tools to efficiency transcribe files and train them on how to use it. We specify a comprehensive set of guidelines for transcription and enforce them strictly. We also use data mining and information retrieval tools to analyze the transcript text and detect out-of-context terms and mistakes.

We have designed this system to produce the best quality transcript in shortest possible time and it has proven itself over time. We have completed more than 149,653.45 minutes of transcription to date.

The following are some unique features which our system provides.

- Accurately measure the transcript progress depending upon what stage it is in. You can check the progress or view the partial transcript anytime.
- Ensure that there are at least two different sets of ears and eyes on each file which results in higher quality transcript.
- Produce consistent results through data analysis tools which detect inconsistencies across your files.

We also provide free re-reviews if the transcript quality is not satisfactory. In case we detect a very poor audio quality file, we refund the complete amount for the order.

# Appendix 4: Call for Interviewees on Docear Blogg.

## **Be a Guinea Pig: Take Part in a User Study and Help Improving Docear MindMapping**

Wednesday, April 20th, 2011

We – that is the Docear team and Hilah Geer (Head librarian at the Oslo School of Management) – need your help!

Hilah is a fan of Docear and for her research she wants to perform a user study (interviews) with some users of Docear MindMapping to understand how academics are using Docear MindMapping. The results will help us to understand how you are using the software and how we can improve it. Therefore, please, if you are an active academic user of Sciplore MindMapping, help us understanding how you work with Docear MindMapping. Participate in the 30minutes interviews and read on what Hilah has to tell you about the interviews:

### **Details about the study**

#### **About me:**

I am been a librarian since the early 90s. . . . I am exploring the field of Science & Technology Studies as a means of getting a better grasp on the role technology plays in my work. I would very much appreciate your help!

#### **About the interviews:**

**Who will be interviewed:** Anyone who is working in an academic/research environment who uses Docear to any extent with any regularity is valuable to my research. I will be interviewing as many of you as possible.

**How will they be conducted:** I personally will be conducting all of the interviews. They will be done either on Skype or via telephone. They will take 30 minutes. I will be more than willing to go more in depth if the interviewee is interested. Otherwise, I will protect your time and stick to 30 minutes.

**What will they include:** first I will want to gather standardized information: demographics, field of study, institution(s), hardware, software, and networked resources. Then I will have a series of questions to steer a discussion of how you use Docear and why.

**When:** The main interview period will be between May 2. and June 24 (There is a degree of flexibility as regards dates). I will be available for interviews on weekends but weekdays are preferable.

**Conditions:** Of course the results of the interviews will be made entirely anonymous. Upon arrangement of the interviews you will be given a signed *terms of agreement* that you can choose to alter if you wish. All processed anonymous data will be shared with Docear, unless requested otherwise.

**Send an email to [REDACTED] to arrange an interview or ask a question about them. And Thanks!**

