The Annotative Practices of Graduate Students: Tensions & Negotiations Fostering an Epistemic Practice

by

Marie-Eve Bélanger

A thesis submitted in conformity with the requirements for the degree of Master of Information Studies
Faculty of Information
University of Toronto

© Copyright by Marie-Eve Bélanger 2010
Released under the WTFPL free software license.
The Annotative Practices of Graduate Students: Tensions & Negotiations Fostering an Epistemic Practice

Marie-Eve Bélanger

Master of Information Studies
Faculty of Information
University of Toronto
2010

Abstract

This research explores the annotation and note-taking practices of graduate students at a large Canadian university and reports on the sets of activities, habits, objects, tools and methods that define the practice. In particular, this empirical study focuses on understanding the integration of annotation and note-taking practices within larger scholarly processes such as reading, analyzing and writing. This study therefore aims to describe and analyze annotation and note-taking not only as the material externalities of the research process, but also as crucial epistemic practices allowing students to progress from one research activity to the other. Interviews are supplemented by document collection and analyzed using a multi-perspectival framework based on theories of practice. The findings describe an annotation lifecycle and suggest a new model of the scholarly process using annotation and note-taking practices as units of analysis. The study further discusses annotation as a primitive epistemic practice and examines the productive tensions and contradictions fostering the student’s progress towards her goals. Addressing annotation practices in light of the rise of digital textuality, this research finally proposes requirements for future tools supporting scholarly practice.
Acknowledgments

This thesis is the result of a long journey that benefitted from the advice, help and guidance of multiple colleagues, peers and friends.

I would like to sincerely thank my supervisor, Matt Ratto, for facilitating this journey. His patience, assistance and commitment were crucial for the completion of this thesis. I thank him for challenging me (in the best way possible!) and teaching me to not be afraid of my own crazy ideas. I am also grateful for the Critical Making resources that Matt made available to me and would like to thank all members of the lab for their continual support and feedback over the last year.

I am also grateful to Alan Galey for his valuable feedback and support throughout this endeavor. Alan brought a different perspective to the table, which created a productive re-assessment of my research. I thank him for asking the right questions at the right time and for his positive attitude throughout this journey.

If not for Matt and Alan, this thesis would have never existed. I feel extremely privileged to have worked alongside such passionate individuals.

I also want to acknowledge the work of my colleagues at the Faculty of Information who took the road less traveled – the thesis option. Alison, Antonio, Armin, Jamon: this thesis would not be the same without our mutual support. More than anything, our lab sessions helped fill an obvious gap in the culture of research at the Faculty of Information, especially at the Master’s level. It is my hope that future iSchool students find a similar environment conducive to research (find your peers!) and get the necessary support from the Faculty.


Merci pour tout!
# Table of Contents

Acknowledgments ........................................................................................................................................ iii

Table of Contents ..................................................................................................................................... iv

List of Tables ........................................................................................................................................ viii

List of Figures .......................................................................................................................................... ix

List of Appendices ................................................................................................................................... xi

1 Introduction .......................................................................................................................................... 1
  1.1 Defining the Problem: Shifting Practices in Light Digital Textuality ..................................... 1
    1.1.1 Background to the Problem ................................................................................................. 1
    1.1.2 Statement of the Problem ...................................................................................................... 3
  1.2 Defining the Research: Situating Epistemic Practices .............................................................. 5
    1.2.1 Research Questions ............................................................................................................... 6
    1.2.2 Scope and Limitations of the Research ............................................................................... 8
  1.3 Outline of the Thesis ......................................................................................................................... 8

2 Review of the Literature ...................................................................................................................... 10
  2.1 Annotation as Visible Traces and Marks: Form, Content and Function of Annotations.. 11
    2.1.1 Anatomy of an Annotation..................................................................................................... 11
    2.1.2 Annotation as an Associative Device .................................................................................. 13
    2.1.3 Annotation as a Visual-Spatial Vocabulary ......................................................................... 17
      2.1.3.1 The Taboo of Annotation ................................................................................................ 18
      2.1.3.2 The Book Versus the Text: Consequences of Annotation .................. 19
    2.1.4 Dimensions of Annotation: Deriving Meaning from Visual Qualities ......................... 20
    2.1.5 Functions of Annotation ........................................................................................................ 21
      2.1.5.1 Higher-Order Functions ................................................................................................. 22
      2.1.5.2 Finer-Grained Functionalities ....................................................................................... 25
    2.1.6 Zooming In: Synthesis of the Section ................................................................................... 28
  2.2 Contexts of Annotation Practice: Personal, Professional, Social and Institutional Influences .... 30
    2.2.1 Annotation and the Scholarship Process .............................................................................. 30
      2.2.1.1 Scholarship & Information Behavior ............................................................................ 31
      2.2.1.2 Practice-Based Studies of Scholarship Process ......................................................... 35
3.3.2.2 Symbolic Markings ................................................................. 100
3.3.2.3 Mixed Markings ...................................................................... 102
3.3.2.4 Color and Typographic Treatment ........................................... 104
3.3.2.5 Encoding and Context: Consequences of Spatial Constraints on Annotation Format .................................................. 105

3.3.3 What are the Preferred Tools and Mediums to Annotate and Take Notes? .......... 106
3.3.3.1 Digital Tools ........................................................................ 107
3.3.3.2 Print-based vs. Screen: The Rise (and Acceptance) of Digital Annotation? .................................................................. 117
3.3.3.3 Physical Tools ....................................................................... 119
3.3.3.4 Devices, Supports, Tools and Diversity ...................................... 121

3.3.4 Annotation and Note-Taking Lifecycle ........................................ 122
3.3.4.1 Trigger .................................................................................. 123
3.3.4.2 Capture ................................................................................ 124
3.3.4.3 Transfer ............................................................................... 127
3.3.4.4 Maintain .............................................................................. 129
3.3.4.5 Refer .................................................................................. 132
3.3.4.6 Complete ............................................................................ 137
3.3.4.7 Discard ............................................................................... 138
3.3.4.8 Archive .............................................................................. 138

3.3.5 How Does the Research Process Affect the Nature of Annotation and Note-Taking? .................................................................. 139
3.3.5.1 Acquainting Activities: searching, retrieving, reading ............. 143
3.3.5.2 Analytical Activities: Reading, Organizing, Interpreting, Outlining ... 144
3.3.5.3 Composition Activities: Drafting, Writing .................................. 146

3.3.6 What are the Attitudes of Students Concerning Their Own Practice? .......... 149

3.4 Discussion .................................................................................. 151
3.4.1 Annotation and Note-Taking as Epistemic Micro-Practices .................. 152
3.4.1.1 Annotation and Note-Taking as Micro-Practices ...................... 153
3.4.1.2 Annotation and Note-Taking as Epistemic Practices ................... 159
3.4.2 Annotation and Note-Taking as Materialities of Infrastructures ........... 168
3.4.2.1 Digital Tools and the Myth of Active Reading ......................... 169
3.4.2.2 Mediums, Formats and File types ............................................ 173
3.4.2.3 Social and Disciplinary Norms ............................................... 176
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.3 Annotation and the Rise of Digital Textuality</td>
<td>178</td>
</tr>
<tr>
<td>4 Conclusion</td>
<td>185</td>
</tr>
<tr>
<td>4.1 Limitations of the Study</td>
<td>190</td>
</tr>
<tr>
<td>4.2 Future Work</td>
<td>191</td>
</tr>
<tr>
<td>References</td>
<td>193</td>
</tr>
<tr>
<td>Appendices</td>
<td>201</td>
</tr>
</tbody>
</table>
List of Tables

Table 1: Types of Association Differentiated by Shape, Proximity and Position on the Page .... 16
Table 2: Seven Dimensions of Annotation by Cathy Marshall ........................................... 21
Table 3: Functionalities of Annotation .................................................................................. 28
Table 4: Comparison of Research Cycle Models Using Chu's (1999) as a Basis ..................... 34
Table 6: Brief Description of the Participants ........................................................................ 74
Table 7: Supplementing O'Hara's Typology of Reading Goals (1996) ................................. 91
Table 8: Comparison of Scholarly and Information Work Primitives ...................................... 156
List of Figures

Figure 1: Anatomy of an Annotation .......................................................... 13
Figure 2: Example of a Composite Association ........................................... 14
Figure 3: Example of a Node-to-Annotation Association ............................... 14
Figure 4: Example of a Standard Association ............................................. 14
Figure 5: Example of a Word-to-Word Association ................................... 15
Figure 6: Vincent Placcius's Note Closet ....................................................... 24
Figure 7: Bookwheels and Their Users ....................................................... 51
Figure 8: Visual Representation of the Different Understandings of Schatzki (2001) & Simpson (2010) ............................................................................. 62
Figure 9: Visual Analogy for the Structure of our Framework ...................... 62
Figure 10: Zooming In (left) and Out (right) on Practice .............................. 70
Figure 11: Situating our Research Objectives on the Zoom In and Out Scale Using the Cube Analogy ........................................................................... 72
Figure 12: Three Examples of Selections Created by Highlighting or Framing Certain Passages ................................................................. 81
Figure 13: Example of a Selection Created by Copying and Pasting Text from the Source File to an External Notes File (the red rectangle is ours) ............................................................................. 81
Figure 14: Example of a Highlight Used as Part of a Larger Annotative Act .... 82
Figure 15: Example of an Emphasis Mark Used Concurrently with Highlighting ................................................................. 83
Figure 16: Example of Comments Appended to a Text .................................. 84
Figure 17: Example of Interpretive Notes in the Form of an Early Project Outline ................................................................. 85
Figure 18: Example of a Project-Related Meta Note in a Calendar ................ 86
Figure 19: Diversity of Meta Notes on a Spectrum Ranging from Project to Content-Related... 86
Figure 20: Typology of Notes and Annotations ........................................... 88
Figure 21: Typology Transposed to a Spectrum ........................................... 88
Figure 22: Example of Restructure and Labeling of a Text ............................ 94
Figure 23: Multiple Types of Markings on the Same Page ............................ 103
Figure 24: Pre-Designed Arrows Packaged with the PDF Reader Software are Used to Point Out Important Content .................................................................... 106
Figure 25: Comparison Between the Toolbars of Adobe Acrobat (top) and Skim (bottom) .... 109
Figure 26: Notes Created Using the Note Tool in Preview (left) and Adobe Acrobat (Right) . 109
Figure 27: Asterisks Created by Placing Single Lines at Different Angles ........ 110
Figure 28: Example of an Intermediary File Collating Notes, Quotes, Thoughts and Other Materials ........................................................................................................... 111
Figure 29: File Types Differ According to Research Activity (Left: Mix of Pages and Text Edit Files for Notes; Right: Mix of Pages and Word Documents for Drafts) ................................................. 112
Figure 30: Source Material and Notes are Often Stored in Close Proximity........................... 114
Figure 31: Comparison of E-Books Interfaces ........................................................................ 115
Figure 32: Highlighted Files .................................................................................................... 121
Figure 33: Annotation and Note-Taking Lifecycle, Adapted from Lin, Lutters & Kim (2004). 123
Figure 34: Shortening of the Cycle due to Time Limitations ............................................... 127
Figure 35: Expanded Refer Stage ......................................................................................... 133
Figure 36: Re-Use Cycle of Annotations and Notes (with Expanded 'Refer' Stage).............. 138
Figure 37: Volume of Meta and Research Notes Created (Top) and Used (Bottom), Throughout the Research Project........................................................................................................... 141
Figure 38: Volume of Reading and Personal Notes (2 Types of Research Notes) Created (Top) and Used (Bottom), Throughout the Research Project ................................................................. 142
Figure 39: Types of Notes and Annotations Observed in the Field....................................... 144
Figure 40: Two lifecycles Emerging in the Early Days of the Project .................................... 144
Figure 41: Dominant Lifecycles During Analytical Activities ............................................. 146
Figure 42: Annotation Lifecycle During the Composition Activities ..................................... 149
Figure 43: Situating Annotation and Note-Taking Practices in the Context of the Research Project ................................................................................................................................. 157
Figure 44: Situating Annotation and Note-Taking Practices in the Context of the Research Project, Mediating the Research Object and its Partial Representations ............................................................. 159
Figure 45: Annotation Lifecycle ............................................................................................ 163
Figure 46: Simplified Representation of 2 Overlapping Social Worlds, and the Different Roles for Common Actors and Objects ........................................................................................................ 177
Figure 47: Spectrum of Annotation Types ............................................................................ 186
Figure 48: Annotation Lifecycle ............................................................................................ 188
List of Appendices

Appendix A: Recruitment Letter for Interviews and Document Collection .................. 201
Appendix B: Interview Guide ...................................................................................... 202
Appendix C: Ethics Approval ...................................................................................... 204
Appendix D: Letter of Consent sent to Participants ......................................................... 205
Appendix E: Table summarizing types of notes and research activities ...................... 207
1 Introduction

1.1 Defining the Problem: Shifting Practices in Light Digital Textuality

1.1.1 Background to the Problem

Research into scholarship and disciplinary practices has steadily contributed to the field of information science. Concurrently, studies of computer-mediated work, digital information appliances and digital scholarship have increased in the fields of human-computer interaction and digital humanities. While these research trends originate from very different scholarly traditions, they however converge in their aim to understand the different facets of scholarly work in light of the recent increase in availability of digital sources and tools. In particular, information practice studies have examined the characteristics of information seeking and use as performed on screen and online, and assessed the shifting sets of scholarly activities and habits. Human-computer interaction studies have attempted to support the scholar’s transition towards the screen by developing requirements for digital scholarly tools. Finally, digital humanities research has attempted to situate the new role of digital material and resources in the scholarly work of humanists.

These studies acknowledge that scholars and students have come to rely on the computer for their daily work. While digital sources and tools are increasingly visible in the scholar’s research process, the extent of use of such resources is still unclear. For instance, the literature on the facilitation and support of analytical reading is divided when it comes to the screen. One perspective on the matter states that the majority of academic users continue to print out digital documents and read them in paper form (Shaikh, 2004; Liu, 2005; Tenopir, King, Edwards & Wu, 2009). These empirical studies claim that academic reading is still dominated by paper. Therefore, while users take advantage of the increased accessibility of search and retrieval that digital textuality affords, scholars and students may consider screens, interfaces and reading appliances problematic for scholarly use and thus tend to default back to print-based habits. However, other studies reveal that sustained screen reading may be indeed migrating to the screen; this may be due to the adaptation and adaptability of current academic users (Nicholas, Rowlands, Clark, Huntington, Jamali, & Ollé, 2008). This contradiction points to two larger
implications. First, scholarly practices are currently shifting and have yet to stabilized: certain practices such as reading and annotating are still oscillating between paper and screen. Second, patterns of use may be extremely localized and influenced by external factors: the tendency of certain groups of users to adopt digital activities may be partly caused by their local disciplinary or academic contexts.

Despite the explicit advantages and innovative possibilities that digital textuality may afford, this contradiction also reveals a certain resistance on the part of the user, not only to the technological shift itself, but also to the changes in behaviors and practices that these innovations impose. Lanzara (2010), echoing Brown & Duguid’s reasoning in *The Social Life of Things* (1996), highlights how long-established practices can be medium-specific, that is “the objects, tools, routines and representations that constitute the practice are specifically dependent upon the medium in which they have been formed and live” (pp. 1369-1370). This may explain the lack of enthusiasm and widespread adoption of reading technologies and devices for scholarly purposes.

While many companies have invested important funds in the development of digital paper lookalikes and tailored LCD displays for greater onscreen readability, most of them failed to account for the multiple corollary practices associated with analytic reading (e.g., annotating and writing) that may themselves evolve when ported over to the screen. These activities foster an active reading strategy (Adler, 1940a) and allow the reader to actively (and materially) engage with the text. Therefore, these devices perpetuate a passive mode of reading by limiting the active construction of meaning-making and fostering a consumptive reading behavior. The passive reading assumption built into these devices may begin to explain why the adoption of screen-based academic reading continues to lag behind the adoption of e-paper or other reading devices dedicated to mainstream materials.

Research into active reading has long been interested in print-based practices, describing and analyzing how academic research unfolds in a written and paper-based environment. The strongest and most fruitful line of research on active reading stems from the human-computer interaction discipline, yielding important design requirements for digital reading devices based on their paper counterparts. Yet, this area of research also assumes a continuity of reading practices (and their corollaries) when transposed from print to digital. Practices tend to be simply
transposed to the screen without fully assessing the potential changes in activities and habits prompted by the digital textuality.

In his essay entitled “Digital Documents and the Future of the Academic Community”, Peter Lyman warns the scholarly community against such assumption (Lyman, 1999). Lyman questions the continuity of scholarly communication, a term designed to encompass the academic research process, its infrastructures, and its outputs in the context of print publications and digital communications. The impact of new technologies in the academic milieu must be understood not only in terms of increased productivity and efficiency, but also in terms of the creation of new information products, leading to a shift in the core organization of academic work and in the overall academic culture. Lyman calls for a serious re-assessment of the consequences of the modes of production that govern knowledge work, and ponders “whether printed knowledge and digital information are used in the same way” (p. 367).

Perhaps, then, it is fair to call for a re-examination of scholarly practices in order to truly assess how practices tend to change according to the mode of production or the medium. Peter Lyman’s call for a shift of focus in the analysis of scholarly activities cannot be ignored. While artifacts and externalities of the scholarly production must be considered, “the focus should [also] be on the function of the artifact and market in creating and sustaining the social worlds, or communities, of the readers.” (p. 374)

1.1.2 Statement of the Problem

The awkward socialization of new technologies, that is, the mutation and shift in meaning and purpose as a new artifact crosses different social contexts, has been well documented and researched from multiple angles (Latour, 1986; Bowker & Star, 1999; Walsham, 2001). This theme has been mirrored in multiple research projects concerning the academic context. Specifically, valuable research has been undertaken concerning the shift to digital libraries (Borgman et al., 2005), the use of e-books (Carden, 2008; Mazza, 2008), and the adoption of learning management software (Tutty & Klein, 2008; Hutchinson, 2007) to name a few.

Worthy of further exploration is an alternative conception of the issue, specifically looking at how current scholarly practices are enacted and performed in an everyday context, shifting the focus from the technology to the activities, methods, contexts and objects of the scholarly
process. While the research process and information activities of scholars have been at the core of multiple research projects in the last few decades (specifically centered on the broad topics of reading, researching and writing), other more discrete scholarly activities have been poorly addressed by research. These discrete activities are the building blocks of significant boundary practices normally ensconced between larger, more recognized academic functions (e.g., reading, researching, writing). These practices, such as annotation and note-taking, are ingrained in scholarly practice. They sit at the intersection of artifacts and social practices as a result of the physical engagement of the scholar with the materials and objects around her. In light of the issues stemming from the rise of digital textuality and the hypotheses that practices are indeed medium-specific, these discrete scholarship practices need to be explored further.

Annotation and note-taking practices have recently garnered sustained interest from the information community for their role as inherent components of the information practices of scholars. These studies have overwhelmingly shaped our current understanding of the scholarship process as a serial, temporally-based system where annotations can be tied to a specific phase of the cycle. While these models of scholarly research shed some light on the scholar’s activities and processes, they however fail to account for the wide variety of annotation practices permeating the various stages of the research cycle. These studies also often employ a very narrow definition of annotation and note-taking and build in different assumptions, reducing this practice only to its material consequences, educational benefits or cognitive demands. Although most of these studies recognize the tight coupling of reading and annotating, the assumptions built into their operational definition of annotation deny the integration of this practice in the larger, more holistic scholarly context. However, recent work by Unsworth (2000) and Palmer & Cragin (2008) hint at the pervasiveness of annotation practices within the scholarship cycle, terming such discrete activities “scholarly primitives”. According to Unsworth, these “basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation” (np) form the basis of every higher-level scholarly activity.

Building on the notion of the scholarly primitive, but distancing itself from the linearity of previous models, this thesis explores the bridging function of annotation and note-taking practices in the ecology of the graduate student. Looking at annotation and note-taking as a modular and recurrent information primitive, this study situates these practices as visible traces
of information access, with variable functions and goals, according to the type of information work involved from ideation to dissemination.

1.2 Defining the Research: Situating Epistemic Practices

The broad aim of this research is to assess how annotation and note-taking practices, as situated intellectual and material processes, are shaped by general scholarly practices, models of the academic community and the infrastructure inherent to the practice. Annotation is defined in our study as any extra information about another piece of information, attached to the original document without modifying it (Hunter, 2009). Jackson (2001) further distinguishes between marginalia (the descriptive or analytical notes made by the reader and found on the margins of the page), from the non-verbal codes and signs of the reader’s attention (such as asterisks, underlines and highlights). Our study assesses both types of markings. For instance, comments and notes written in the margin of a book and highlighted passages of a journal article are two types of annotations that are considered in this thesis. Additionally, the definition of annotation used for this study does not discriminate between medium. Hence, annotations created on paper using pens and highlighters are assessed alongside the notes and highlights created in a PDF file. Furthermore, this research takes into consideration the process of creation and use of these marks. Annotation is therefore understood as a process and a mark.

This definition is also exclusively concerned with readers’ annotations. Hence, although the history of authorial and printed annotations (the comments of authors and editors printed alongside the main text [Tribble, 1993; 1997; Connors, 1998]) is rich and may warrant further research in information science, this type of annotation falls outside of the research scope. This research also avoids domain-specific definitions that, while related to this current description, do not fall within the scope of this project. These disciplines and projects include computational biology and the genome annotation project, as well as computer science and the role of documentation in code.

Jackson (2001) also clearly calls for a distinction between notes in books and notes in notebooks. This study therefore considers note-taking separately from annotations. While some of these notes may have similar content and purposes as annotations, their storage on an external medium (such as a notebook) may influence the practice much more differently than annotations. Therefore, our study considers research-related notes not taken on the source material itself:
jottings on scraps of paper, comments in a text file, or passage and quotes copied and pasted into a word processing file.

Switching gears, this study addresses the concept of practice as “embodied, materially mediated arrays of human activity centrally organized around shared practical understanding” (Schatzki, 2001, p. 11). As pointed out by many contemporary authors (Schatzki, Simpson and Bernstein for instance), the term practice is somewhat vague and ambiguous, a buzzword of sorts, whose defining boundaries are fluid enough to permit multiple interpretations and significations, and “imprecise and open enough to allow people from different traditions to join without renouncing their worldviews” (Miettinen, Smara-Fredericks, Yanow, 2010, p. 1313). Despite this vagueness and the lack of agreement on the extent of the concept of practice, this study settles on a concrete definition of practice as sets of methods, habits, tools, objects, activities, and the relationship created between these items (Nicolini, 2010). Hence, this thesis considers annotation in a larger context. Annotation and its inherent activities, tools and objects are examined in light of the reading and writing processes that give rise to the annotation practice.

Concretely, this study aims to fulfill three objectives. First, this research attempts to understand how annotation and note-taking function as epistemic practices in the daily work of researchers and students. Second, scholarly annotations are analyzed as “materialities of infrastructure” (Star, 2002), and thus this study discusses how and why the “invisible layers of control and access” are shaping (and shaped by) these practices. Finally, these practices are examined in light of the recent rise of digital textuality in order to assess if they remain stable as new modes of production emerge or if they shift and modulate in order to adapt to new technological, material and digital environments.

1.2.1 Research Questions

The broad aim of assessing how annotation and note-taking practices are enacted and performed by members of the academic community is complex and multi-dimensional. The literature relevant to answering this question emerges from multiple fields of research, such as human-computer interaction, information needs and use, education and pedagogy, and the history of the books and readers. Upon reviewing and analyzing the literature from these different areas of inquiry, a number of questions were formed and refined. These clusters of questions guide our data collection and analysis and are designed to fulfill our three research objectives.
1. The first set of research questions describes and analyzes the practice and its structure: What are the activities, methods, and objects inherent to these practices? How do they structure the practice? Why are they used, and in what contexts? What are the functions and purposes of annotation?

2. The second line of questioning pertains to the integration of this practice in larger practices and processes: What happens to these practices throughout the research project, from ideation to dissemination? How are they integrated with reading, thinking, researching and writing practices? How is the current academic research model influencing the practice? How can we describe annotation as a materiality of infrastructure? What are the systems of standardization and classification involved in this practice? How are social and disciplinary norms affecting annotation and note-taking practices?


In order to address these questions, an empirical study has been conducted with seven participants. The study centers on the annotation and note-taking practices of graduate students in humanities and social sciences faculties at a large Canadian university. Interviews and document collection are used in order to gain significant insights on how and why such practices are enacted and performed.

Furthermore, a multi-perspectival framework is used to frame our results and analysis. This framework allows the researcher to zoom in and out of practice, addressing the finer-grained details of the materiality of the practice as well as their consequences on the larger scholarly processes. Different theoretical lenses are used in order to gain access to various levels of granularity.

Knorr Cetina’s theory of epistemic practice (1981; 2001) and its focus on partial representations of the research object and its inherent dynamic allows us to zoom in on the doings of our participants, the active role of material elements and tools, the influence of the medium on the
practice, and the methods and strategies deployed by our participants. Zooming out of the
practice entails the analysis of the wider repercussions of the annotation practice on the overall
scholarly activities and the research project. Theories of pragmatism and symbolic interactionism
are used to look at the bigger picture and to trace relationships between the zoomed in and
zoomed out perspectives.

1.2.2 Scope and Limitations of the Research

Contrary to the majority of studies on annotation that opt to focus on specific facets of the
practice (e.g., the materiality, the methods, or the purposes of annotation), this research addresses
the features of the practice holistically. These features of annotation and note-taking are
intrinsically linked to each other and to wider processes framing the practice such as reading,
thinking and writing. This yields a rather large set of variables to study. In order to assess these
variables comprehensively, the scope of the research is restrained to the study of a small sample
of participants. Participants were selected from social sciences and humanities faculties at a large
Canadian university. While this small sample allows the researcher to focus and address precise
annotation practices, it however entails that findings cannot be generalized beyond this small
group.

This research is also restrained to the study of annotations strictly made to textual documents and
to typed or written notes. Hence, this study deliberately omits audio and video documents, and
avoids larger practices involving the annotation of domain-specific material outside of social
sciences and humanities. Furthermore, this study is restrained to the individual research projects
of the participants, therefore only superficially addressing annotation activities in other academic
contexts such as group work and class work.

1.3 Outline of the Thesis

The subsequent discussion covers the following topics. The first chapter reviews the relevant
literature culled from five academic disciplines: the history of the book and reader response
theory, human-computer interaction (HCI), humanities computing (HC), education and
pedagogy, and information needs and uses. This section first zooms in on annotations by
examining their material characteristics and their inherent visual-spatial vocabulary, their content
and dimensions, and finally their functions as derived from their material and conceptual
characteristics. The literature review then zooms out and considers the wider professional, social, intellectual and institutional contexts of annotation. More specifically, annotation and note-taking activities are contextualized within the scholarship process, examined in light of larger processes such as reading and writing, and finally related to the communities and institutions supporting them.

Chapter 3 describes the empirical study undertaken as part of this thesis. The multi-perspectival framework is described and contextualized within the different traditions of theories of practice. The different theoretical lenses used as part of the framework are then examined and related to each other as we explain the steps involved in the zooming in and out of practice. The sampling procedure and research instruments are then addressed. We detail the characteristics of our participants and introduce our interview and document collection strategies. Data analysis, ethical considerations and the limitations of this methodology are then considered before closing this section.

The results and analysis of this empirical study are then revealed concurrently. Here again, this section first zooms in on the content and encoding of annotation and considers the different tools used by our participants. Zooming out, we expand on the annotation lifecycle as unearthed in the field, looking closely at its variations according to different scholarly activities. The annotation practice of our participants then becomes the unit of analysis as we reveal the different phases of the scholarly process as framed by annotation activities. Discussion of these results closes Chapter 3 and carefully addresses the three goals of this thesis as set out in the introduction.

Chapter 4 concludes this thesis and reviews the general findings and insights from this study. After disclosing the limitations of this thesis, we propose new requirements for future tools supporting scholarly practice and point to areas of research that warrants further research.
2 Review of the Literature

This chapter reviews the relevant literature pertaining to annotation and note-taking practices. We first zoom in on annotation by reviewing studies addressing the marking itself: its form, content and functionality. We then zoom out to understand the wider context of annotation creation, organization and use, therefore reviewing studies that situate this practice at a personal, professional, social and institutional level. The materiality of annotation, and its consequences for both levels of analysis, is a dominant thread throughout this discussion.

The literature is culled from five different areas of scholarship: the history of the book and reader response theory, human-computer interaction (HCI), humanities computing (HC), education and pedagogy, and information needs and uses. These fields are selected for their focus on the materiality of annotation and on the contextual understanding of the practice. The history of the book, HCI and humanities computing are mainly concerned with the materiality of the practice, but for different purposes. Scholars of the history of the book and of the reader understand annotations as traces of past readings. Annotations are important traces necessary to “‘historicize’ the conditions of the encounter between the world of the text—which is always a world of forms, supports and objects—and the world of the reader—who is always a reader socially defined by the competency, conventions, expectations and practices of reading that he shares with others” (Chartier & Friedman, 1997, p. 10). HCI and humanities computing have examine the visual-spatial characteristics of annotations and their meaning for the user: defining these visual features and situating them in the reading and scholarship practices is crucial for these fields as they apprehend the shift to digital textuality and deplore the inefficiency of their tools for the digital realm. While most of the HCI and humanities computing studies reviewed for this thesis give requirements and recommendations for new tools and devices supporting intellectual work, most link these findings back to a more conceptual understanding of the matter, resonating with information needs and uses studies. The information studies reviewed in this thesis mainly situate annotation practices as part of a larger cycle of scholarship, define their use for scholars and explore how the information stored in these annotations is integrated in the larger research process. This literature crosses paths more than once with educational and pedagogical studies explaining how intellectual work should be performed.
2.1 Annotation as Visible Traces and Marks: Form, Content and Function of Annotations

As pieces of information materially added to a document or stored externally, annotations in the context of written communication possess a number of visual-spatial characteristics from which users derive meaning. The following section addresses several dimensions of annotations resulting from the material, visual and symbolic aspects of the marking.

2.1.1 Anatomy of an Annotation

Cathy Marshall, researcher at Microsoft Research, is often credited for the first modern in-depth studies of annotations in the field of human-computer interaction. Her numerous studies often address the form and functionality of paper-based annotations. Through empirical research, Marshall and her colleagues assess the materiality of paper annotations in order to recreate such practices in a digital environment, especially via the use of reading appliances.

Her large study of annotated textbooks (Marshall, 1998) reveals the gamut of notation strategies of students. Highlights, underlines, marginal bars or brackets, symbols (such as asterisks), arrows, short phrases or extensive marginalia are all part of the student’s toolbox. In order to address this diversity, Marshall (1998, 2009) and Agosti, Ferro, Frommholz & Thiel (2004) establish a consistent terminology of annotation, aimed at the development of systems and mechanisms for creating, storing and using annotations in a digital context.

Marshall expands on this anatomy in her book entitled Reading and Writing the Electronic Book (2009), building on the work of Agosti, Ferro, Frommholz and Thiel (2004) on the integration of annotation in digital library systems. It should be noted that these studies contain some very specific assumptions tainting their results: annotations are seen as an output of the reading process alone, with little consideration for the future usage of such information. Furthermore, they assess only in situ annotations created using a writing tool, leaving out such practices as bookmarking or annotating using removable media (bookmarks or sticky notes), physical annotations using the affordances of the support material (such as dog-earing) or the practice of taking notes in a separate location (such as in a notebook).

According to this terminology, annotations are generally composed of three basic elements: a body, an anchor and a marker. The body of an annotation refers to the content that the user adds
to the source material, such as a lengthy comment in the margins or an idiosyncratic marking on the text (e.g., an asterisk or a proofreader symbol). In the case of a standalone, highlighted portion of text, the annotation has no explicit body. Marshall explains that these annotations are a common case of *null-content*. However, we should note that highlights and underlines might contain a very specific meaning or functionality and give an implicit content to the mark. The functionalities and purposes of such markings are further explored in section 2.1.4 *Dimensions of Annotations* and 2.1.5 *Functions of Annotations*.

The anchor element defines the scope of the annotation and reveals the link between the source document and the added content. Anchors may be explicit (highlights and underlines spanning a precise passage of text) or implicit (a note written in the general area of a portion of the text, without any visual selection of the content to which it is referring). They may also be broad in scope (a note written on the first page of a journal article reviewing the entire content of the article) or narrow (proofreading marks related to a specific).

The third element of an annotation is typically the marker. This element indicates how the anchor should be rendered when displayed. For instance, the marker of an underlined passage is the colored line under the selected text, the visual characteristics of the underline itself. While Marshall phrases this last element in a very specific manner in the context of the digital display of annotation, this item is crucial in a print-based context. We further discuss the function of the marker in section 2.1.3 *Annotation as a Visual-Spatial Vocabulary*. Figure 1 illustrates the various parts of the anatomy of annotation: its body is the marginal text, the anchor is the underlined source text and the marker is the black underline.
This taxonomy hints at the inherent associative function of annotation. There are multiple ways to create an association between the annotation and the text. These associations often result in different purposes, functionalities and meanings for the user. The following section surveys the visual-spatial characteristics of annotations as an associative device.

2.1.2 Annotation as an Associative Device

The body-anchor-marker trio revealed in the last section can be combined in multiple ways and to different effect. The arrangement of these items produces different types of association, which ultimately influence the meaning of the marks for the reader (see section 2.1.5 Functions of Annotation). Moreover, the material characteristics and the context of these elements (i.e., their shape, their proximity to elements of the document, and their position on the page) affect the type of association created. Marshall (1998) examines four kinds of associations

1) *Composite/Collection*: notes referring to an entire article or chapter (figure 2);

2) *Node-to-annotation links*: comments or symbols localized within a document part but not visibly anchored (figure 3);

3) *Standard association*: comments or symbols visually anchored to a portion of the text (figure 4);

4) *Word-to-word association*: words or shorts phrases written in direct proximity to specific words in the text (figure 5).
KeyStrokes: Personalizing Typed Text with Visualization

Petra Neumann, Annie Tai, Torre Zisk, and Sherliagh Carpendale

Department of Computer Science, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada T2N 1N4

Figure 2: Example of a Composite Association

Figure 3: Example of a Node-to-Annotation Association

Figure 4: Example of a Standard Association
Three mechanisms govern these associations. First, arrows and other lines are used to connect an anchored element in the document (an underlined or highlighted phrase) to its body. Second, brackets and braces, while themselves acting as both a link and an anchor, associate a commentary with the selected text. Finally, proximity is used as a mechanism for association, connecting markings to the text.

We supplement this analysis by including a collapsed association: some standalone markings such as highlights or underlines can effectively collapse the body-anchor-marker trilogy on itself. A standalone highlight acts as the anchor, body and marker altogether, referring uniquely to the selected highlighted word or passage itself without being linked to any other markings. The same type of association operates for symbolic reader’s marks such as asterisks, stars or other idiosyncratic markings. In fact, these symbolic markings are extremely versatile and may carry more meaning for the reader than highlights and underlines. In his book *Used books: marking readers in Renaissance England*, William Sherman (2008) explores the role of these marks in the reading strategies of historic readers. He first notes how the concurrent use of multiple symbols and letters in the marginalia of medieval and Renaissance readers appears to serve an indexing function, simultaneously classifying the information contained in the text and pointing out the interesting passages. Hence, these markings effectively collapse the body-anchor-marker trio.

First, the symbol acts as an anchor: readers place these marks in the margin next to significant phrases or passages. However, the visual characteristics (the marker) of the symbol are generally not insignificant: these characteristics give meaning to the mark and constitute the body of these annotations. For instance, Sherman describes the annotation strategy of an anonymous reader of the 1569’s edition of Cicero’s *De Oratore* and states how the reader…

```
[...] tried out various symbols but once he had settled on a system he wrote it neatly toward the bottom of the page and surrounded it by a rough decorative border: he lists the signs used for tagging passages on particular topics (a trident
```
was used for passages of argumentation or reasoning and the symbol for Venus signaled an interest in love, and so on), and for marking particular rhetorical devices (such as "amp[licatio], " "metap[hor], " and "sim[ile]," each of which is signified by a symbol that looks like a flower). (Sherman, 2008, p. 27)

Sherman then launches into an exploration of the manicule, these pointing hands drawn or printed in the margins of books (e.g., ☞). Manicules operate in a manner similar to other reader’s marks. However, Sherman asserts “that after a signature and a monogram, the manicule was the most personal symbol a reader could develop and deploy” (Sherman, 2005, p. 20). Hence, for the early modern reader, manicules represented far more than a simple mark. They symbolized their marks and “must have played an important role in the personal process of making a book meaningful” (p. 21).

This body-anchor-marker framework therefore allows us to see that careful selection by the reader of visual characteristics such as shape, proximity and location on the page is an important part of the annotation strategies. Table 1 clarifies the importance of visual-spatial characteristics and demarcates the different associations.

Table 1: Types of Association Differentiated by Shape, Proximity and Position on the Page

<table>
<thead>
<tr>
<th>Type of association</th>
<th>Shape</th>
<th>Proximity</th>
<th>Position on the page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite (Marshall, 1998)</td>
<td>Written or symbolic</td>
<td>Greater distance from text indicate a general note</td>
<td>Top or bottom margins, front or endleaves for books (Adler, 1940b)</td>
</tr>
<tr>
<td>Node-to-annotation (Marshall, 1998)</td>
<td>Written or symbolic</td>
<td>Proximity creates the link between the comment and the text</td>
<td>Margins, gutters</td>
</tr>
<tr>
<td>Standard (Marshall, 1998)</td>
<td>Body: Written or symbolic. Link: line or arrow Anchor: thick line (highlight) or thin line (underline)</td>
<td>Proximity is enhanced by the use of a line or arrow as a link. Anchor (highlight) is in the text (distance: 0)</td>
<td>Margins, top or bottom margins</td>
</tr>
<tr>
<td>Word-to-word (Marshall, 1998)</td>
<td>Written</td>
<td>Link created by the very short distance between note and text</td>
<td>Between the lines of text</td>
</tr>
<tr>
<td>Collapsed</td>
<td>Thick line (highlight) or thin line (underline), symbol</td>
<td>Null distance</td>
<td>On the text, in the margin</td>
</tr>
</tbody>
</table>
2.1.3 Annotation as a Visual-Spatial Vocabulary

Other visual characteristics of annotation, such as color and medium\(^1\) are also used to create specific meanings and are significant elements in the visual-spatial system of annotation.

The contrasting color of highlights and other marks is the variable responsible for effective emphasis and beneficial annotation strategies for undergraduates according to Fowler and Barker (1974). Several pedagogues in the education field including J. Wesley Miller recommend using this visual characteristic to distinguish between different topics (Miller, 1980). This echoes the earlier recommendations of 16th and 17th century scholars who endorsed the use of color to flag different types of passage as a means to cope with information overload (Blair, 2003). However, color-coded highlighting strategies appear to be rare for contemporary readers, probably because of the number of tools necessary to create such scheme and of the cognitive load involved in switching tools (Marshall, 1997). Annotation is mainly an unselfconscious activity (Marshall & Brush, 2004; Marshall, Price, Golovchinsky, & Schilit, 1999) and therefore, the act of putting down a highlighter pen of one color to grab the next color interrupts the broader reading task.

The use of different tools to mark up pages and the quality of the marks they create on the page are tangentially related to the benefits of the color characteristic. Certain pedagogues suggest to new analytic readers the use of different tools specifically for their texture. Again, J. Wesley Miller recommends to students to experiment with different writing implements. He proposes a hierarchy of writing mediums: texts should first be underlined using a number 2 pencil, using ink in a second pass to underline key points only (p. 575). Color, texture and durability are different elements at stake when selecting the right marking medium.

The examination of the visual characteristics of annotation opens the discussion to the interesting practices of dog-earning (Blair, 2003, 2004b; Schott, 2007; Daston, 2007), cutting and pasting (Blair, 2003; Sherman, 2008) and marking books with fingernails (Blair, 2003). These practices do not use any intermediary devices (such as pens or highlighters) to create their markings. Rather, they use the materiality of the book and play with the affordances of paper. Whereas

\(^1\) Medium here refers to the material means used to create an annotation, namely the tools such as pen, pencil, highlighters, independently of the medium of the practice itself (paper or digital). It also relates to the “marker” element of annotation as described by Marshall.
more visual methods such as commenting and highlighting are described as a separate layer of the document (O’Hara & Sellen, 1997; Miller 1980), these more physical and tactile practices affect the core of the document itself and alter the physical integrity of the book.

Two themes emerge from the role of visual-spatial characteristics of annotations seen in this section. First, annotation is perceived as a transgression and a taboo. This negative perception is linked to material qualities of the markings and their institutional repercussion. Second, the material characteristics of these markings necessitate that we distinguish between the book and the text and that we address the consequences of these marks on either of these layers. The following sub-sections discuss these two themes.

2.1.3.1 The Taboo of Annotation

First, the rise of the negative perception of annotation (some considering it as a type of book abuse) relates to the visual and tactile qualities of these marks, perceived as a mutilation of the book. These personal notations are described as “messy” (Jackson, 2001, p. 235) and as a “crime against the book” (p. 74). The rise of this negative attitude correlates with the growth of the public library system, which imposes the superiority of the general good to the wishes and personal habits of the individual (Jackson, 2001; Hauptman, 2008). The personal, intimate and visible nature of these markings conflict with the rise of property rights and “proprietariness” that shape the policies of circulating libraries.

This attitude is still present and is written in current library policies, although slightly adapted for digital content. For instance, this excerpt from the University of Toronto’s Library Conduct Regulation compares annotation to mutilation and explicitly prohibits this activity: “Mutilation of library materials or files by marking, underlining, removing pages or portions of pages, removing binding or staples, removing security devices, tampering electronically, or in any other way damaging or defacing library materials” (University of Toronto Libraries, 2010).

The negative perception of annotation troubles many pedagogues and authors advising students to integrate annotation to their reading and studying strategies. Their thoughts on the issue are often conflicted. They praise annotation as a means to fully owning and understanding a book, while carefully warning the students to never write in books that are not theirs. The American philosopher Mortimer Adler perhaps expresses both concerns most dramatically. On one hand,
Adler (1940a) contends that “marking up a book is not an act of mutilation but of love” and that “[f]ull ownership comes only when you have made it a part of yourself, and the best way to make yourself a part of it is by writing in it” (np). However, he clearly understands the institutional ramifications of marking up books and advises students to be respectful of other’s property: “You shouldn’t mark up a book which isn’t yours. Librarians (or your friends) who lend you books expect you to keep them clean, and you should. If you decide that I am right about the usefulness of marking books, you will have to buy them” (np).

However, this institutional perspective on annotation is merely the tip of the iceberg. Readers are not solely concerned with the mutilating and messy aspects of the marks. Rather, most are also displeased by the imposition of others’ readings onto theirs (Jackson, 2001, p. 243). Perhaps Virginia Woolf said it best: “Marginalia that were full of meaning for the original owners of the books are often of little or no interest to later ones.” (as cited in Jackson, 2001, p. 237) The debate of authority and authoritative voice is thus at the crux of this perception, with the voice of the previous reader effectively competing with the voice of the text, to the detriment of the current reader.

2.1.3.2 The Book Versus the Text: Consequences of Annotation

Second, this negative perception brings to light some of the reasons explaining why certain individuals hesitate to digitally annotate documents using standard, multi-purposes reading and writing software such as text editors. In their study of the differences between reading and annotating on paper and on screen, Kenton O’Hara and Abigail Sellen (1997) reveal that this reluctance to annotate on screen is caused by the reader’s perception that “emboldening, italicizing or underlining all alter the original document. Subjects indicated that they wanted to regard annotations as a separate layer of the document, and felt uncomfortable not maintaining this distinction” (O'Hara & Sellen, 1997, p. 137).

This example, as well as the previous discussion concerning the negative perception of annotation, point to a certain confusion between the text and the medium supporting the text. The dominating discourse here is one where the material qualities of the object (printed or digital) are subsumed into the text. Therefore, affecting or modifying the material qualities of the book is also affecting the text. While the distinction between the text and the medium warrants further exploration, it however falls outside of the scope of this thesis. Section 2.2.2 Annotation in
Reading and Writing will however address these issues tangentially in light of annotation and note-taking practices.

Annotation is thus much more than a simple trace left by the reader on a page: it is a visual-spatial system combining features such as color, shape, proximity, location and medium. Due to the meaning-making nature of annotation, its visual characteristics may best be defined and described in the context of the content and function of annotation.

2.1.4 Dimensions of Annotation: Deriving Meaning from Visual Qualities

As a visual-spatial vocabulary, annotation may carry different meanings. The previous sections reveal the various components constituting annotations and explore the different associations created between these items. Hence, multiple variables affect the meaning of annotations: the anatomy of annotation, the types of association, and the visual-spatial qualities of the situated marks. The current section then turns to the content and dimensions of annotation to further explain the meaning-making system at play in annotation practice.

Ovsiannikov, Arbib and McNeill (1999) provide one of the most explicit and focused discussion of annotation content in their article entitled “Annotation Technology”. Following their wide review of the literature, they define annotation content as “the semantics carried by the marking” (p. 337). This implies that annotations can be semantically explicit (and thus understandable by others than the annotator herself) or idiosyncratic and very personal in meaning. Also at play here is the notion of semantic distance between the annotation and the original text to which it refers. This semantic distance can be null (such as in the case of a highlighted word) or have a high value (such as doodles or remarks completely unrelated to the original text). Semantic distance is however not necessarily correlated to the semantic explicitness of the annotation: a note in the margin can be clear and spell out precise thoughts, but can be completely unrelated to the text under study.

Cathy Marshall examines these two attributes at length in her seminal study “Toward an ecology of hypertext annotation” (1998) terming them “dimensions of annotations”. Marshall’s description of seven dimensions highlights the importance of a contextual understanding of the content of an annotation. It should be noted that these dimensions were devised with the specific intent of shaping design requirements for a subsequent reading device. The characteristics
discussed in this scheme are thus very explicit and often tied to general patterns of the practice that can be easily integrated in a digital environment. The seven dimensions devised by Marshall are reproduced in Table 2.

**Table 2: Seven Dimensions of Annotation by Cathy Marshall**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
</tr>
</thead>
</table>
| Formal v. informal                             | **Formal:** Structured, standard metadata, often used to ensure interoperability.  
|                                                | **Informal:** Less structured notes, including reader’s symbols              |
| Explicit v. tacit                              | **Explicit:** Related to intelligibility, could be read and understood by others than the annotator.  
|                                                | **Tacit:** Has meaning for the annotator alone                               |
| Annotation as writing v. annotation as reading | **Annotation as writing:** wordy, lengthy explanations, use of marginalia, closer to the “postmodern ideal of the polyvocal text, the text as a participatory medium in which annotators are writers” (Marshall, 1998, 42).  
|                                                | **Annotation as reading:** attention symbols and structure, organization of content |
| Hyperextensive v. extensive v. intensive        | Annotations that refers to and links with other texts (hyperextensive, extensive), or that refers to itself (intensive) |
| Permanent v. transient                         | **Permanent:** annotation may have value beyond its first creation and use (i.e., an annotation created in a library book may have value for a subsequent reader).  
|                                                | **Transient:** annotation has limited value and usefulness, ephemeral.        |
| Published v. private                           | **Published:** closer to authorial and published annotation  
|                                                | **Private:** personal annotations                                            |
| Global v. institutional v. workgroup v. personal | **The public of the annotation: who is it created for and who will benefit from it?** |

### 2.1.5 Functions of Annotation

As many researchers discovered, meaning is often hard to derive from a few marks on the page. This might be the reason why most historians of the book studying reader’s annotations are careful about radical assumptions in their findings. Based on the literature reviewed, and also from our own lived experience, it is clear that annotation systems are idiosyncratic. However, common tools and shared visual-spatial vocabulary may reveal common usage and functions in reader’s annotation.

Annotation is often considered an externality of the reading process and the material consequence of goal-oriented reading and thinking. Hence, researchers can gain an understanding of these larger scholarly processes by studying the visual qualities and content of annotations. Studies in HCI and in the history of the book often examine at the material qualities
of the markings to derive the potential functionalities of annotation for the reader. While many studies point to the unselfconscious nature of this activity, it is clear that readers engage with the text in a number of ways and for a variety of purposes, often without clearly knowing their own intent when doing so (Marshall, Price, Golovchinsky, 1999).

The main objective in annotating texts, however, is to set apart important information, regardless of whether the markings are created consciously by the reader or not. Functions of annotation are generally divided in two categories: higher-order functions and finer-grained functionalities. The subsequent sections address these two perspectives. Higher-order functions tend to be more general and address larger processes at play between the annotative act and the student. Conversely, finer-grained functionalities are derived from smaller processes and are often a direct result from the material characteristics of annotation. The next sections examine both categories and emphasize the functions necessary for scholarly, research-based annotations.

2.1.5.1 Higher-Order Functions

Higher-order functions address the general modes of interaction between the individual and the annotation. Whereas finer-grained functionalities are often closer to the materiality of annotation, and perhaps more easily translated into design requirements for annotation tools, higher-order elements define larger categories of use and point to the interconnections between annotation and other scholarly practices.

In their survey of computer science students and professors, Ovsiannikov, Arbib & McNeill (1999) discuss four functions of annotation linked to four wider processes. Annotations are created and used to remember, think, clarify and share. As a means to remember, the authors observe that the visual contrast between the marking and the source document helps draw the attention of the reader and eases the task of re-finding information at a later time. An annotated document can be quickly searched for highlighted keywords in order to remember the text and its main points. Scholars of the history of the book also point to note-taking as a means to remember. Blair’s work on the Jesuit scholar Jeremias Drexel indicate that note-taking and excerpting were the preferred ways to memorize material, therefore supplementing the reader’s own memory (Blair, 2004b, p. 98).
Annotations are also created to help the reader think critically. This is normally true of marginalia including critical remarks, questions and notes written in the margin. Hence, annotations are an easy way to store new ideas prompted by the text with relatively minimal cognitive overhead. A third function exposed by Ovsiannikov, Arbib & McNeill is the use of annotation to clarify certain ideas in the text. The authors note that, contrary to annotations used for thinking, these clarifying notes do not carry new information and ideas. The benefits of these annotations are not necessarily immediate. They rather often involve a future usage: “[r]ephrased in the person’s own conceptual language, ideas are much easier to work with, which saves the reader time later when reading the paper again” (p. 336). The final function uncovered by the authors does not specifically stem from their survey results. It is rather an hypothesis to be tested in their annotation framework: annotation is used for comment sharing. This function is thought to be more frequent in the review process when the annotator must communicate information related to the source document to her collaborators.

Ann Blair, a scholar specializing in the cultural and intellectual history of early modern Europe, obtains similar results and discusses similar functions. In her paper entitled “Note Taking as an Art of Transmission” (2004b), she identifies storing, sorting, summarizing and selecting as the four basic “maneuvers” involved in annotating or note-taking.

Blair understands storing in a larger historical context. Storing is not solely related to the information concretely contained in marginalia and notes stored in books or notebooks. Rather, she understands the concept inclusively: the methods of storage play an important role in the transmission of knowledge to this day, despite their specific constraints of reliability, preservability, and accessibility. This unavoidably directs our attention to the materiality of the storage medium, whether the source document itself or any external media. It furthermore points to the material consequences of storing these notes in a wider system: Blair highlights the use of the note closet (see figure 6), a device used for storing, sorting and, perhaps more intriguingly, to foster collaboration. Indeed, this closet was designed to ease collaboration between members of a literary society or within a group of students. The closet allowed for easy pooling of notes and consisted of “dozens of moveable slats labeled with topical headings, which swivel to access note slips for each headings on which notes are available.” (p. 105)
Storage invariably implies a certain sorting system for ease of retention and retrieval. While this sorting function was performed by integrating the information to be sorted into narratives in oral cultures, written cultures have devised various types of ordering: alphabetically, systematically, or miscellaneously. Commonly used for reference books in the Renaissance (and often noticed in the arrangement of computer files), the latter type accounts for the creation of finding aids and index.

Blair then observes how summarizing and selecting are often explicitly intertwined: “Note taking differs from the transmission of whole texts in that only parts of a whole are selected for transmission. The note taker can process many texts in this way and can integrate the selections from different sources into one set of references.” (p. 86) Selecting parts of a larger text entails the paraphrasing or summarizing of passages of the source text. Blair describes two note-taking methods common to the 16th and 17th century, and recommended by Francis Bacon: “by Epitome, or Abridgement or under Heads and Common Places” (Bacon, cited in Blair, 2004b, 86). The former operates by reduction: summarizing the original text and presenting the notes in
the form of *adversaria*, in the order of the source text. The latter operates by selection: the reader selects passages of the original text for their ideas and styles, and presents them under thematic or topical headings, forming a *commonplace*. The role of the commonplace, a notebook collating selected passages from multiple sources, will be further explored in section 2.2.3.2 *Annotations and Institutions*.

### 2.1.5.2 Finer-Grained Functionalities

Studies discussing finer-grained functionalities attempt to look beyond higher-order elements and describe a more detailed understanding of how annotations operate concretely in the field. It is not surprising, therefore, that a number of these studies ultimately attempt to derive functional requirements for annotation tools and devices by examining finer-grained functions arising from the materiality of reading.

Cathy Marshall (1998) reconstructs a considerable number of these functions by studying a wide sample of textbooks. Her results point to the metacognitive aspect of the traces left by readers. The first functionality she surveys is the use of annotation as a *procedural signal*, distinguishing between what is important for a certain task and what is not. This may be indicated by highlighted higher-level items (such as section headers), or conversely by crossed out paragraphs. As *placemarks and aids to memory*, annotations are used to store away potentially important information for future usage: emphasis symbols such as asterisks characterize such functionality. Third, annotations seem to be an *in situ way of working problems*. Explanations and notations found beside equations often demonstrate this function. We should note that this function is highly context sensitive and perhaps a consequence of her textbook sample. Marshall then identifies extended highlighting as a way of *working through a complex narrative*. Fifth, marginalia and in-text commentaries often function as a *record of the interpretive activity of the reader*. Finally, she considers markings that may simply *be incidental*: doodles, notes to self and drawings, however unrelated to the reading, often find their way into textbooks.

While her findings have been widely cited, it is important to put these results in perspective. First, her study does not address the use of removable or external media (e.g., notebooks or bookmarks). Second, this research uses mainly textbooks as her object of study. Consequently, her results pertain to one type of literature only. Annotation of other text genres such as scientific journal articles is ignored. Additionally, her study does not address the perspective of the reader:
interviewing students and readers could have contextualized annotation further, by giving the opportunity to the main subjects to reflect on their markings. While these criticisms relate to issues that may be outside of her research scope, it nevertheless points to an unspoken assumption permeating her study. Marshall addresses annotation practice as a unidimensional process and therefore ties this practice to a strict conception of learning and studying. Studying needs to be done “by the book”, and is indeed restricted by the textbook. Discounting the practices that may occur in notebooks or on removable media therefore restricts our understanding of annotation to a practice that is solely oriented towards reading and always performed in situ.

Other studies suggest that the material context of the reading and annotating practices influences the functionalities of annotation. Whereas Marshall’s study generates findings restricted to the paper-based textbook, other researchers examine how a change in the support medium (shifting from paper-based to digital practices) influences the function of annotations. Agosti, Ferro, Frommholz & Thiel (2004) assess the collaborative aspects of annotation in the context of a digital library system. Their overarching finding echoes multiple hypertext studies of annotation (O’Hara, Taylor, Newman & Sellen, 2002; Schillit, Golovchinsky & Price, 1998; Liao, Guimbretière & Hinckley, 2005): the linking functionality of annotation is predominant on screen and in interactive interfaces. This functionality is salient in research on reading devices. Xlibris, one of the early pen and tablet reading devices out of FX Palo Alto, allows for hyperlinks to be created by users between different documents and even automatically generates sets of suggested hyperlinks to other documents and sources (Schillit, Golovchinsky & Price, 1998).

In the digital library context, users create annotations that link multiple documents, creating alternative browsing paths for themselves and for others using private or public annotations accordingly. Interaction, access and sharing are three finer-grained functionalities derived from the linking affordances of the digital environment (Agosti, Ferro, Frommholz, & Thiel, 2004). Nested annotations (i.e., where users can react to other users’ comments and notes) support a strong collaboration, earmark of the interaction functionality. Conversely, a weak collaboration occurs when an individual can only view annotations created by others, without being allowed to comment. This supports the access function of annotations in a digital library system, where
users rely on alternative browsing paths created by other users. Finally, the sharing function occurs when users share their personal thoughts and ideas with a wider community.

This study builds on a specific context and understanding of annotation. In this digital system, annotations are conceptualized as a layer sitting on top of the source material. This layer is comprised of three (or more) potential sub-levels related to the privacy of the note: a private layer accessible only by the annotator, a collective layer shared by a pre-determined group of people and a collective layer for the use of the larger digital library community. The functionalities mentioned in these studies do not strictly arise from the perspective of annotation as an output of the reading process. Rather, annotations are here perceived as a communication act not necessarily tied to the reading process. Hence, these functionalities involve other types of scholarship practices, and take into account the sociality of annotation-as-collaboration.

Kenton O’Hara’s research agenda while at EuroPARC revolved around a similar idea and addressed annotation in contexts other than reading. O’Hara specifically examines the goal-oriented intellectual work involved in working from multiple source documents, often culminating in the composition of scholarly texts. In his 1996 technical report for Xerox, O’Hara discusses the functionalities of annotations and notes in goal-oriented reading and writing from multiple sources. He describes annotation as an “intermediate text that functions as a means for planning, reviewing, and re-shaping ideas and information in the source texts and the reader’s own thoughts.” (p. 11) O’Hara’s research seems to be unfortunately unique in its area: our literature review yields very few studies addressing the role of annotations in phases of the scholarship cycle other than reading. O’Hara’s research is also significant for its inclusive definition of annotation. His technical review also include annotation and notes on supports and materials other than on the source material itself: O’Hara carefully reviews and analyzes the characteristics and functionalities of notebooks and the uses of intermediate documents such as drafts and outlines. We expand on the notion of annotating while writing in section 2.2.2.3 Ways of Writing as part of the contextual understanding of annotation.

Table 3 summarizes the functionalities of annotation reviewed in this section. We indicate the material and intellectual contexts of these studies to emphasize the context-dependent qualities of annotations and their functionalities. These variables, as well as other factors such as the
institutional, intellectual and social contexts, shape the annotation practice performed by individuals. These factors are directly addressed in the next section.

### Table 3: Functionalities of Annotation

<table>
<thead>
<tr>
<th>Author</th>
<th>Material Context</th>
<th>Intellectual context</th>
<th>Functionalities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In situ annotation</td>
<td></td>
<td>• Placemark and aid to memory;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• In situ way of working problems;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Record of interpretive activity;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Incidental</td>
</tr>
<tr>
<td>Ovsiannikov, Arbib, &amp; McNeill (1999)</td>
<td>Paper-based</td>
<td>Reading process</td>
<td>• To remember;</td>
</tr>
<tr>
<td></td>
<td>In situ annotation</td>
<td></td>
<td>• To think;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• To clarify;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• To share</td>
</tr>
<tr>
<td></td>
<td>In situ annotations</td>
<td></td>
<td>• Interaction;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Access;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Sharing</td>
</tr>
<tr>
<td>O’Hara (1996)</td>
<td>Paper-based</td>
<td>Reading from multiple</td>
<td>• Planning;</td>
</tr>
<tr>
<td></td>
<td>In situ annotation</td>
<td>sources Writing process</td>
<td>• Reviewing;</td>
</tr>
<tr>
<td></td>
<td>Removable media</td>
<td></td>
<td>• Re-shaping ideas</td>
</tr>
<tr>
<td></td>
<td>Notebooks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blair (2004)</td>
<td>Paper-based</td>
<td>Reading process</td>
<td>• Storing;</td>
</tr>
<tr>
<td></td>
<td>Removable media</td>
<td></td>
<td>• Sorting;</td>
</tr>
<tr>
<td></td>
<td>Notebooks</td>
<td>Writing process</td>
<td>• Summarizing;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Selecting</td>
</tr>
</tbody>
</table>

### 2.1.6 Zooming In: Synthesis of the Section

This section zooms in on annotation by first considering the materiality, visual-spatial qualities and functions of annotation and note-taking. Studies addressing the materiality of printed annotation often stem from human-computer interaction disciplines. These studies tend to reveal general patterns of written annotation that can be reproduced digitally. While annotation and note-taking strategies tend to be idiosyncratic and greatly diverse, common items and general
patterns of association exist. Hence, the anatomy of annotation as described by Marshall (2009) describes different constituents always present (even if collapsed) in annotations. This anatomy also points to the associative function of annotation (i.e., relating the anchor to the content – or body – of the annotation). Marshall (1998) examines four different types of associations: composite, node-to-annotation, standard and word-to-word. We add to these categories the “collapsed association” most commonly used for standalone symbolic markings such as highlights or reader’s marks. These types of associations are crucial for our study as they may carry different purposes and meanings for our participants.

Color and texture are visual qualities that also affect the functions and purposes of these different associations as they relate to the “marker” component of the anatomy of annotation (i.e. color-coding strategies may refine the meaning of different associations, especially for the collapsed association). These visual-spatial qualities of annotation are the root of the negative perception of annotation that stem from the first days of the circulating libraries and that is perpetuated to this day. In this institutional context, annotation is perceived as a mutilation of the book and in fact blurs the boundary between the book and the text. Indeed, the negative perception of annotation subsumes the material qualities of the book into the text: affecting the book is also altering the text. This points to the great impact of visual-spatial qualities of annotation on other aspects of the practice. Annotation is thus much more than a simple trace left by the reader on a page: it is a visual-spatial system combining features such as color, shape, proximity, location and medium.

In fact, these visual characteristics are key to the different dimensions that annotations carry. Hence, the single use of reader’s marks or highlights denote an informal and tacit dimension of annotation, whereas extensive marginalia can be more explicit and formal. Selecting the proper visual qualities of the mark is key for the reader: different types of mark have different functions. Our review of the literature yielded two broad categories of functions: higher-order functions characterized by their broad understanding of the interactions between the reader and the text, and finer-grained functionalities, which tend to be closer to the materiality of the mark and may vary extensively depending on the qualities of the support medium.
2.2 Contexts of Annotation Practice: Personal, Professional, Social and Institutional Influences

Annotation practice is difficult to observe and analyze as a standalone entity. It is intrinsically linked to the activities of reading and writing: annotation, as a written mark, typically occurs as a response to working closely with all types of materials (e.g., books, articles or drafts). Moreover, a single annotating event (i.e., the moment of creation, management or use of annotation) may happen at different phases of the research process. To this effect, research normally associates annotation with the reading and analyzing phases of the scholarly process (Adler, 1940a; 1940b; Marshall, 1997; 1998; 2009). However, much evidence points to the presence of annotation towards the later phases of the research process, such as drafting and writing (O’Hara, 1996; O’Hara & Sellen, 1997; O’Hara, Taylor, Newman & Sellen, 2002). The typical research cycle seems to be influenced by wider factors such as the academic community and academic disciplines, the introduction of new technologies, and the publishing industry. Each of these factors tends to affect not only the overall research cycle, but also the practices, habits and activities constituting the larger phases of the cycle. This section reviews the larger context of scholarly annotation to reveal the interconnectedness of the multiple practices involved in this ecology.

2.2.1 Annotation and the Scholarship Process

Scientists and their research process have been at the core of multiple information-seeking studies since the early days of the information field in the 1950s. The general assumption of these early studies is a very functional one: the research process can be brought to light by studying the scientist’s journal and library usage, citation patterns and communication (Case, 1986). While research on the information-seeking habits of humanists and social scientists has been markedly less abundant than studies of natural scientists, large research projects, for the most part originating in the sixties, have shed some light on the processes and habits of these scholars.

These previous studies, as well as methodological improvements, allowed for a behavioral and practice-based turn in the literature to emerge in the eighties. These two trends (behavioral and practice-based) are further explored in this section as they depart from a mechanistic view of the user and move towards a more grounded and contextual understanding of the practices of
scholars. This section thus reviews the literature on information needs and uses of humanities and social sciences scholars, situating annotation either as a discrete activity, a process, or a *scholarly primitive* in context.

### 2.2.1.1 Scholarship & Information Behavior

Departing from a tradition of studies concerned with requirements analysis, information behavior studies are generally centered on the information needs, seeking and use behaviors of individuals (Wilson, 2000). This shift also indicates a change in methodologies used, from quantitative to qualitative methods, as the field focuses its attention on the individual rather than the system. Thus, subsumed under the heading of information behavior are distinct and finer-grained fields of studies such as information seeking, information needs and information use studies.

While research in this field was at first concerned with understanding the demographic composition of library clienteles and their needs, it quickly moved towards the exploration of the information behaviors of bounded communities, using scientists and scholars as recurring subject of studies. This is due to the nature of scholarly work, where the “information seeking activities of researchers take place within the context of the research process and are linked with the various stages or phases of that process” (Chu, 1999, p. 250). While the literature contains some descriptive studies, most of the research surveyed attempts to model the information behavior of researchers, often referring to *research-phases, features or stages*.

#### 2.2.1.1.1 Descriptive Studies: Disciplinary Comparisons & Early Understandings

Weintraub’s exploration of the different modes of scholarship of humanists can be considered as a foundational text (Weintraub, 1979). While still very library-centric, Weintraub attempts to characterize the intellectual and material habits and needs of the scholar. Comparing the humanities work processes to the physical sciences, the strength of this paper can be found in the discrete characteristics of humanistic work that Weintraub reveals in passing, thereby igniting a trail of further studies. Most notably, he recognizes the influence of environmental and institutional factors on the work practices of the scholar, qualifies the humanist’s topic of interest as “open-ended”, and links this characteristic to a non-sequential workflow. Fabian (1986) echoes these two latter points, hinting at the constant refinement and readjustment process underlying the scholar’s work. This dynamic emerges from the interplay between the
researcher’s questions and the replies obtained from the sources (p. 83). This description of the research process implicitly suggests various and complex ways of reading. Fabian recognizes at least two primordial types of reading (reading from source materials and reading from the state of research), but doesn’t probe any further to assess how and why these activities are undertaken.

2.2.1.1.2 Modeling Research: Abstraction & Processes

A second and perhaps stronger trend in information behavior research focuses on modeling information practices and abstracts the primordial phases of the research process. This shift is based on the assumption that variations in information behavior can be linked to stages or phases of scholarship (White, 1975). While the overarching field of human information behavior yielded multiple generic models (Wilson 1999; 2000), the following analysis features models tailored to scholarly research and employs research features rooted in academic research as their basis. Hence, these process models “allow us to grasp the breadth of practices that require information support and the nuances of their disciplinary culture and requirements” (Palmer & Cragin, 2008, p. 185).

Research Phases Models

A first approach to modeling the research process of scholars is to decompose the process into groupings of related tasks constituting phases or stages. An early model of the information gathering behavior of academic economists (White, 1975) contains three stages (problem, methodology and presentation). This macro view of the research process is the result of a “combination of functions in each stage and the differences in behavior patterns associated with each function [which] explain the variations in the communications behavior in the research stages” (p. 341). Thus, to each stage is associated a different information function such as 1) perception of problem 2) definition of problem 3) development of methodology 4) provision of data 5) suggestion of information source 6) analytical assistance and 7) practical assistance. Interestingly, White does not conceptually define and describe each of these functions: how is a problem perceived? What are the cues used by the researcher? While the research phases proposed by White seem linear, she also posits that the distinction between the stages are artificially imposed, suggesting that a researcher may be overlapping multiple phases or projects. She proposes that these stages should be understood as “a measure of progress into a research project” (p. 344), ultimately giving in to the linearity of her model.
Sue Stone (1980; 1982) refines the three stages brought forth by White and proposes a meso level model based on her empirical studies. Her model acknowledges outside influences on the researcher (phase 1: thinking and talking to people) and then separates the remaining phases between reading (2: reading what has been done in the field, 3: studying original sources and making notes) and writing (4: drafting the write-up and 5: revising the final draft). Similarly to Fabian, she implicitly acknowledges different types of reading and furthermore points to the role of annotation within the research cycle (phase 3). Her model can be considered as a meeting place for both intellectual and material practices of the researcher. This is somewhat echoed by Chu’s recent model (1999) who gives much more emphasis to the notion of analysis and dissemination. Chu’s model also explicitly acknowledges the cyclical nature of research with her 6th and final stage further writing and dissemination, implying the re-work of ideas into new final products.

Case’s 1991 empirical study involving twenty American historians alludes to even shorter micro-cycles within a single research process. While stating that research-stages are illusory, Case uses a processual terminology to name various categories of work. Case therefore points to the active nature of the researcher when choosing and refining topics, planning and conducting studies, gathering and interpreting evidence, and writing and revising manuscript. He explicitly states that these clusters of tasks may go on concurrently, despite the inherent linearity implied by the names of these categories. The author understands the overarching research process as one of category refinement, originating with an open-ended quest slowly refined into concrete categories “via a complex process we call ‘writing’” (p. 63). Case thus points to the importance of note-taking and annotation practices in the doing stage, shifting to proper academic writing in the writing phase. The refinement of categories, as well as this push-and-pull between intellectual and material practices creates the “back-and-forth nature of this kind of scholarship” (p. 77). Table 4 summarizes and compares the various models assessed.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Idea</td>
<td>1. Thinking and talking to</td>
<td>1. Problem</td>
<td>Choosing and refining</td>
</tr>
<tr>
<td></td>
<td>people</td>
<td></td>
<td>topics</td>
</tr>
<tr>
<td>2. Preparation</td>
<td>2. Reading what has already</td>
<td>2. Methodology</td>
<td>Planning and conducting</td>
</tr>
<tr>
<td></td>
<td>been done in the field</td>
<td></td>
<td>studies</td>
</tr>
<tr>
<td></td>
<td>3. Studying original</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>sources and taking notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Elaboration</td>
<td></td>
<td></td>
<td>Gathering and interpreting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>evidence</td>
</tr>
<tr>
<td>4. Analysis and</td>
<td>4. Drafting write-up</td>
<td></td>
<td>Writing and revising</td>
</tr>
<tr>
<td>writing</td>
<td>5. Revising the final draft</td>
<td></td>
<td>manuscript</td>
</tr>
<tr>
<td>5. Dissemination</td>
<td></td>
<td>3. Presentation</td>
<td></td>
</tr>
<tr>
<td>6. Further writing and dissemination</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2.1.1.3 Limitations of Research Models and Research-Phases and Implications for Annotation Research

While the surveyed models give us an appreciation of the breadth of higher level processes involved in the research cycle, the very nature of modeling - constantly abstracting from reality and situated practice - may limit our understanding of annotation practices. A second and more pressing limitation of research models may be the silences left around the interstices between the various phases and stages. A number of questions surfaces: how and why does a researcher move from one phase to the other? What are the activities involved in the progression from one stage to the next? The notion of category refinement as advanced by Case (1991) failed to fully explicate the research progress.

While most of the authors previously surveyed agree that the notion of stages or phases is illusory (Case 1991; White 1975) and that multiple stages can happen concurrently within a single project or across multiple projects, the very nature of the phases elicited by these authors implies the notion of temporality and sequentiality (which is carried over in the naming convention of certain phases such as *starting* or *dissemination*). Furthermore, as pointed out by
White, the concept of research-phases builds in the assumption of a clear separation and distinction between the various tasks of the researcher. While these assumptions are productive to refine our understanding of the overall research process, perhaps a different approach may bring to bear different activities or facets of the research process. This may be especially fruitful for situating annotation and note-taking within the scholarly cycle. Hence, we posit that in order to deepen our understanding of annotation as a scholarly activity, we must assess the different disciplinary, social and institutional contexts giving rise to the scholarly cycle.

Information behavior literature, both descriptive and model-based studies, seems to have a very narrow view of annotation in the context of the research process, either mentioning it only in passing, or even failing to mention such activity. Hence, according to most of the information behavior literature surveyed (White, 1975; Case, 1991; Chu, 1999), annotating and note-taking cannot be considered as a research-phase. This is explained in several ways. First, higher-level models fail to capture finer-grained, material activities such as annotation, giving precedence to larger intellectual processes (as seen in White, 1975 and Chu, 1999). Second, annotating and note-taking activities are often folded into other phases of the research process (as seen in Stone, 1980). Finally, annotating and note-taking may not completely fit the description of a research-phase bounded by a separation of tasks and a certain time span. Consequently, this points to the more pervasive nature of such activity. When explicitly mentioned, such as in Stone (1980) and Case (1991), the authors acknowledge the dual nature of the activity, both intellectual and material. Case also alludes to a certain transitional function of annotation, helping the researcher “absorb” information (p. 76).

2.2.1.2 Practice-Based Studies of Scholarship Process

Perhaps even more fruitful for this study is the literature originating from the recent practice turn in social sciences (Palmer & Cragin, 2008). Lamenting information-centric approaches for their narrow view of the real world of scholarly work, Frohmann (2004) advocates for a practice-based framework in order to understand the multiple activities defining the research process. Palmer and Cragin (2008), citing Schatzki, argue that the term practice, not unlike information, is generally imprecisely used in the literature, which gives rise to a multitude of practice-based approaches. For the purpose of this study and the literature review at hand, we borrow Palmer & Cragin’s reworking of Schatzki’s definition of practice, consisting of arrays of human activities
and habit that are “materially mediated” and “organized around shared practical understandings” (Palmer and Cragin 2008, p. 169 citing Schatzki, 2001). This definition recognizes the importance of addressing the why and how of practice and prioritizes the material and social aspects of information work. Hence, this approach emphasizes not only the roles of tools and resources in supporting and mediating the scholar’s activities, but also the social dimension of the practice, linking the individual to its social and institutional contexts.

2.2.1.2.1 Practice-Based Studies of the Research Process

Similarly to the tradition of information behavior studies, practice-based studies of the research process first dwelled on the figure of the scientist working in natural sciences. The choice of natural sciences as the setting for the first practice-based studies of research work may be linked to the obvious materiality of this research practice, especially when bounded by a location such as the laboratory (Latour & Woolgar, 1986). Conversely, addressing the materiality of humanities and social sciences may be problematic. The scholarly practices of humanists or social scientists are not neatly bounded by a laboratory, nor are they explicitly material like the work of scientists observing proteins or DNA sequences (Knorr Cetina, 2001).

Practice-based studies of humanists and social scientists revolve around the use of materials in settings such as libraries, or focus on their use of documents and texts in larger processes such as reading or writing. The work of Brockman, Neumann, Palmer and Tidline (2001) as well as Palmer and Neumann (2002) is crucial to our understanding of humanist practices. Departing from the traditional metaphor of the library as laboratory, both studies address the day-to-day practices of researchers, based on large-scale qualitative studies of respectively 33 and 25 participants.

Commissioned by the Council on Library and Information Resources in Washington, D.C., Brockman, Neumann, Palmer and Tidline assess the practices of humanities scholars in light of recent technological developments. They identify four types of scholarly practices - reading, networking, researching, and writing - and decompose each of these practices into core activities. Note-taking and annotation practices appear to be intertwined with every type of scholarly practices. They furthermore allude to the varying function, format, and role of annotation depending on the type of practice deployed by the scholar. The creation, organization, and use of annotation then seem to be closely following the research process. Their findings remain
grounded in the materiality of practices and reveal the idiosyncrasies and difficulties of such intellectual work. They specifically single out the writing process, which appears to be complex and demanding even for the seasoned scholar.

Palmer and Neumann (2002) set out their study of interdisciplinary humanities scholars with the broad task of surveying “...the many elements that affect how research is carried out, on the individual, day-to-day level and within the larger work setting that includes institutions, organizations, and researchers’ domains of interest” (p. 87). Building on Case’s (1991) study, they identify the disciplinary norms and the institutional academic context as two influences on the work practices of scholars. The authors identify clear features of humanities research work: 1) while humanities scholars work from a core selection of resources, their path of inquiry is unpredictable, 2) multiple types of reading, often intertwined with the writing process, are practiced by the scholar throughout the project and, 3) technologies are adopted if they fit into established patterns of research (p. 98). The last two findings are especially appropriate for our topic and are examined in their own section (see 2.2.2 Annotation in Reading and Writing).

Both studies prompted subsequent research into work processes and activities. The work of John Bradley in the field digital humanities is especially worth noting here for its contribution to practice-based studies of the research cycle. Bradley’s work (Bradley 2005; 2008; Bradley & Vetch 2006) addresses the role of annotation in the analysis process of the scholar, based on the findings from Brockmann, Neumann, Palmer & Tidline and Palmer & Neumann. He proposes a tool, Pliny, to support conventional scholarly interpretation, surprisingly distancing himself from the current trend in digital humanities to propose tools transforming scholarly practice. As he grounds his software in practice-based research, he also brings forth the important metaphor of annotation, note-taking and writing as portal mechanisms, allowing the public sphere (i.e., the previous literature) to be brought back into the personal worksphere of the scholar and vice versa (i.e., allowing the dissemination of results back into the public sphere).

2.2.1.2.2 Scholarly Primitives

Emerging from the practice-based literature is the notion of the scholarly primitive, first brought forth by John Unsworth (2000). These “basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation” (np) form the basis of every higher-level scholarly activity. Unsworth identifies annotating as one of the primitives, among
his list otherwise comprised of discovering, comparing, referring, sampling, illustrating and representing. As Palmer and Cragin (2008) rightfully pointed out, very few studies fully address scholarly work at this level of granularity, generally painting scholarly activities in broad strokes in the course of widely scoped research (p. 185). Palmer and Cragin define even finer grained primitives termed *information work primitives* including chaining, annotating and browsing. These primitives are normally associated with the larger information work processes of reading, collecting, searching and writing. They therefore distance themselves from Unsworth’s generalizing view and suggest that these primitives should not be extracted and generalized across domains or disciplines. Rather, and this is especially true for designers of information systems, they should be used as starting points for customization of distinct scholarly applications.

This recent work on scholarly and information work primitives is reminiscent of the concept of *information-seeking features* as described by Ellis (1989; 1993). Ellis argues that these features (*starting, chaining, browsing, differentiating, monitoring, extracting, verifying and ending*) interact and interrelate differently according to the unique situation of the researcher, pointing to the idiosyncratic nature of the individual’s research cycle (Ellis 1989, Wilson 1999). These features conceptually sit between the notion of primitive and research-phases. Indeed, features such as *starting* and *ending* point to a certain sequentiality of features in the research work of scholars, while the granularity of the activities brings them closer to the notion of primitive.

### 2.2.1.2.3 Implications for Annotations

The practice-based studies of the scholarly process reveal the adaptable qualities of annotation, a practice seemingly intertwined with or, at least occurring concurrently with other major practices such as reading or writing. Indeed, the practice of annotation greatly varies throughout the research process. This practice is affected by the dominant research-phase at the moment of creation or use of the annotation. Hence, annotations (both the material qualities of the marks as well as the process of creation or use) may be strikingly different whether the scholar is in the early stages of her research (reading around) or finishing up (drafting and writing).
2.2.2 Annotation in Reading and Writing

The previous section highlighted the importance of reading and writing for scholars in their work routines. It is important here to distinguish between reading and writing as phases of the scholarship process (or, rather, an accentuation of the reading and writing activities for specific periods of time), and reading and writing as tasks and activities performed by scholars for a variety of purposes. This section surveys the current research on reading and writing as practices, including typologies of reading and their implications for annotation, the materiality of reading and writing, and the shift from paper-based to digital reading and writing.

2.2.2.1 Ways of Reading (On Paper)

Kenton O’Hara undertakes perhaps one of the more thorough investigations of reading as a scholarly practice in his report entitled “Towards a Typology of Reading goals” (1996). His research aims to “deepen our understanding of how the affordances of paper support the task of reading and comprehending documents” (p. 4). O’Hara posits that readers have a variety of reading strategies at their disposal and that the appropriate strategy is determined by the reader’s goals and motivations. Annotation and note-taking are a support activities that may be tailored to different reading styles. These support activities help the reader achieve his reading goals within the navigational and manipulation constraints of the medium. O’Hara mentions that the reader normally selects the appropriate support activity (i.e., underlining, note-taking, outlining or networking) according to the selected reading style and the ultimate reading goal.

He distinguishes between four reading styles, based on Lunzer’s previous research on reading (1979): receptive, reflective, skim and scanning. According to his typology, a text can be read in a linear fashion or in an incomplete manner (serial vs. non-serial reading) and can be read once or repeatedly. These reading styles imply a different use of annotation and note-taking: a receptive reading (a style approximating listening behavior) will often yield only attentional marks such as underlines, while a reflective reading (a style where reading is often interrupted by reflective moments) may foster the use of note-taking and outlining.

Reading goals also influence the choice of a specific support activity. O’Hara lists twelve reading goals: to learn, to self inform, to search/answer question, for research, to summarize, for discussion, proof-reading, to write and revise documents, for critical review, to apply, for
problem solving and for enjoyment. This typology hints at the pervasiveness of reading throughout the research cycle (e.g. to write and revise documents). This is demonstrated by our mapping of O’Hara’s reading goals to the research-phases model of Chu (1999) as seen in table 5. As with reading styles, different reading goals necessitate different support activities: reading to write and revise document may entail outlining and note-taking, while reading to search/answer question may only produce highlights and underlines. Consequently, this research provides two important aspects for annotation: 1) annotation and note-taking practices are performed throughout the research cycle, 2) different reading styles and reading goals generate different annotating and note-taking strategies.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Idea</td>
<td>Reading to learn</td>
</tr>
<tr>
<td>Preparation</td>
<td>Reading for research</td>
</tr>
<tr>
<td></td>
<td>Reading to summarize</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Reading for critical review</td>
</tr>
<tr>
<td>Analysis and Writing</td>
<td>Reading while writing from multiple sources</td>
</tr>
<tr>
<td></td>
<td>Reading to search / to answer questions</td>
</tr>
<tr>
<td></td>
<td>Proof-reading</td>
</tr>
<tr>
<td></td>
<td>Reading for text revision</td>
</tr>
<tr>
<td>Dissemination</td>
<td></td>
</tr>
<tr>
<td>Further Writing and Dissemination</td>
<td></td>
</tr>
</tbody>
</table>

Other studies refine and expand O’Hara’s typology by looking at specific disciplines and areas of scholarship. Brockman, Neumann, Palmer and Tidline (2001) expand O’Hara’s classification by tailoring the typology of reading to the humanist’s needs and purposes for reading. Findings indicate that scholars do background reading, comprehensive reading and continual reading. They read around a person or period and relate to primary materials by reading for details and to become immersed. Here again, the authors understand reading and annotating as intimately connected. They however depart from previous studies of reading by highlighting how the type and format of annotation vary according to the type of source material tackled. Indeed, primary
source materials often necessitate different annotation strategies due to their content (i.e., primary materials are analyzed differently than secondary materials) and their format. The latter highlights the impact of materiality on the choice of annotation strategy. For instance, the humanist’s primary material (e.g., books, documents) may often be fragile or only briefly accessible through the services of a librarian. Consequently, scholar must use removable media to take notes, or, are required to photocopy the source in order to highlight or write comments.

Many authors address the role of annotation in the natural sciences (Knorr Cetina, 2001, Latour & Woolgar, 1986; Rheinberger, 1997; 2003). While these markings fall outside of our research scope, Rheinberger’s insights concerning epistemic things and the materiality of research may be beneficial for our own study (see section 3.1.2.3 Objectual and Epistemic Practice for a discussion of epistemic objects). According to him, research scribblings are different from any other types of annotation or note-taking strategies. He states that they are closer to the “the materialities of scientific work than are research communications, insofar as the scribbles are quasi parts of the research objects, and therefore have a share in what will become the passive voice of those objects” (1997, p. 314).

Palmer and Neumann (2002) identify three distinct modes of reading for interdisciplinary humanities scholars (scanning, rereading, and reading for writing). The latter two generate a different type of written record of the intellectual work. According to the authors, rereading is facilitated by note-taking. Annotations and notes guide further reading and are conversely refined by subsequent reading. This is especially true of reading for writing, when original sources are reread. In this instance, notes are created to “fix the intellectual work of reading in a primitive form for future development” (p. 100). The annotations and notes created when reading for writing are therefore more than the documentation of an idea. They are formative in themselves, helping the scholar see the relationships between concepts and ideas.

Perhaps a key concept that has garnered sustained attention in this last decade is the notion of “active reading”. Originally described by Mortimer Adler in 1940, the notion of active reading opposes the prevalent assumption that reading is by nature a passive activity fueled by consumptive behaviors. Rather, Adler discusses the dynamic nature of reading and implores readers to actively engage with the text by annotating it: “reading, if it is active, is thinking, and
thinking tends to express itself in words, spoken or written. The marked book is usually the thought-through book.” (Adler, 1940a).

Active reading recently re-emerged in HCI studies of reading devices and interfaces. Active reading in most cases (Schilit, Golovchinski & Price, 1998; Kopak & Chiang, 2009) is the key concept justifying the integration of reader’s tools in screen-based reading interfaces designed for scholars. While some of these devices and interfaces eventually proved unsuccessful, the researchers nevertheless succeeded in building on Adler’s work by refining and augmenting the concept of active reading. Recent research has shown that active reading is highly goal-oriented (Kopak & Chiang, 2009; Wolfe, 2000; Marshall et al., 1999; Marshall and Brush, 2004), gives rise to a “critical engagement” (i.e., the “interplay between information as encountered and the analysis and use of that information” [Kopak & Chiang, 2009, p. 115]), and also involves chaining, chasing, and finding related material (Schilit, Golovchinski & Price, 1998).

2.2.2.2 Ways of Reading (On Screen)

The characteristics and process of reading from screen garnered sustained interest in the last few decades. Research compared the affordances of paper vs. screen, often yielding long lists of user requirements for adapting paper-based reading onto the screen. However, other studies explored reading on screen as a practice separate from reading on paper, with its own sets of activities and objects. This section reviews the differences between paper and screen-based practices and delves into the practice of reading on screen, looking at the specificities affecting annotation and note-taking strategies.

Analyzing the activity of reading, as well as comparing the affordances of different reading interfaces, is not an easy task. Comparative studies often tackle reading with great difficulty: they use constraining methodologies and have wildly variable definitions of “reading”. These nonetheless foundational studies can be divided between process measures and outcome measures. The latter focuses on what the reader ultimately takes away from the text (and her efficiency in doing so), while the former situates reading in a wider context, looking at the uses of the text for the reader (Dillon, 1992).

Outcome measures long dominated the field of reading research, but were recently discredited for their lack of understanding of the context of reading. These studies survey the differences in
speed, accuracy, fatigue, comprehension and reader’s preference between paper and screen\(^2\). However, many of these studies conclude that the differences between reading on paper and reading from screen are “unremarkable or inconsistent” (O’Hara & Sellen, 1997, p. 335). This failure to find any considerable differences is linked to the measures used in these studies (i.e., assigning unusual tasks to participants), the restrictive methodology and sample, and the rapid evolution of new technologies. The latter is especially true for a key finding of this era: reading from screen tends to be slower than reading from paper. This finding was challenged many times: the technological and technical improvements of screens, as well as the adaptation and adaptability of contemporary readers, recently closed the gap between print-based and digital reading.

Early process studies also refine our understanding of the wider reading context and shed light on the affordances that distinguish paper from screen. Manipulation and navigation differences are at the core of these findings. Manipulating documents involves a certain manual dexterity (e.g., holding a book, turning pages, using a finger as a location aid) that can’t be replicated with electronic text (Mangen, 2008). Screen reading however requires the manipulation of mouse and scrollbars, impacting the reading flow of the user. We should note that the studies mainly address the activity of reading from a traditional computer screen. It would be wrong to subsume all types of screens (and types of reading from screens) under the same descriptor. For instance, the smartphone screen is very different than the iPad or the laptop screens. Additionally, while we recognize that multitouch devices may make fruitful use of paper metaphor in their interaction scenarios (e.g., turning a page, using a finger as a location index), these studies fell outside of the broader scope of our research. This interesting question would definitely warrant further exploration and probing.

O’Hara and Sellen (1997) directly link the failure of previous studies to uncover any significant difference between paper and screen to the measures used. They echo Dillon’s concern (1992) that a misconception of the nature of reading often taints the methodology and that researchers “seem to concern themselves with the control of so many variables that the resulting

---

\(^2\) See Dillon, 1992 for a thorough review of this type of literature, which, even if almost 20 years old, still provide valuable insights.
experimental task bears little resemblance to the activities most of us routinely perform as ‘reading’” (1997, p. 335). Their approach then involves an experimental task closer to the everyday activities of their participants: summarizing from multiple texts. Their findings unveil major differences in support activities such as annotation and note-taking. Paper affords multiple types of annotation to be made quickly, concurrently with reading. This is predominantly linked to the materiality of the document and to the interplay between multiple documents (or multiple pages of a single document), which may be displayed and rearranged quickly on a work surface.

Marking on paper allows the reader to extract the structure of the text while providing the wider context for the marking. Conversely, taking notes away from the text (e.g., in a notebook) has a purpose of its own, which is to re-structure ideas and collate information from multiple sources. These notes normally take the form of outlines or plans, continuously modified during the course of reading and writing. The authors observe that even when the notes are taken away from the original source, the process is done without disrupting the main reading task, due to the spatiality afforded by paper (i.e., the spatial layout of documents on a work surface).

Conversely, reading online documents (in a browser) appears to be very different, especially when considering annotation and note-taking. Graphic elements inherent to the reading interface such as toolbars and scrollbars tend to hinder smooth reading and note-taking. Marking the document itself seems to be complex, if not impossible for many test subjects: the limitations and inflexibility of interaction techniques via mouse and keyboards disrupt the flow of reading. Moreover, annotating the original source (if the text was editable) is perceived as altering the integrity of the document. This differs from annotations on paper, perceived as an additional of layer on top of the original document.

The restricted field of view caused by the limited real estate of the screen is one of the main constraints of reading on screen. In their study, O’Hara and Sellen found that, since only a page or part of a page is displayed at once on a computer screen, participants tend to feel lost and that “much of the necessary contextual information for developing a sense of text and location lay beyond the window boundaries” (p. 340). Participants have to make a conscious choice when selecting and displaying specific documents since, unlike an actual workspace where one can lay out multiple document pages (Marshall, 1998; Ovsiannikov, Arbib & McNeill, 1999; O’Hara, Taylor, Newman & Sellen, 2002), the computer screen denies any periphery.
Issues of materiality and interaction are at the core of recent comparative studies (Mangen, 2008; Hiilesund, 2010). In her research on hypertext fiction reading, Anne Mangen (2008) investigates the influence of haptics on the sense of immersion a reader obtains when reading a text. She observes how different types of materiality impact our embodied reading experience differently. She claims that it is the intangibility and the volatility of digital reading that concretely affects the reading experience and fosters a shallow reading style. The materiality of the reading appliance (e.g., book or a screen) limits the scope of possible interactions. Materiality consequently affects the distance between the text and the reader and the location of the interaction. According to Mangen, interaction with digital texts “is experienced as taking place at an indeterminate distance from the actual text, whereas when reading print text we are physically and phenomenologically (and literally) in touch with the material substrate of the text itself” (p. 405). This distance created by the materiality of the text and of its reading appliance directly shapes the annotation practices of the reader. Hiilesund reiterates the importance of bodily and material aspects of reading in his study describing various digital reading spaces: reading “in addition to decoding and comprehension, is also physical handling of reading appliances and that the way we read is dependent on technologies, implement design and text composition.” (np).

These arguments for the reconsideration of the bodily and material aspects of reading are reminiscent of the early discussions on the conventions and orders of the book brought to the fore by French historian and historiographer Roger Chartier and his lineage (Blair 2003; 2004a; 2004;, Topham, 2004). According to Chartier (1994), texts cannot be separated from their support or their materiality: the structure of the text supports a specific delivery format. Hence, two mechanisms operate in the creation of meaning for the reader: the strategies of writing and the author’s intentions, as well as the formal qualities of the book, often the result of a long series of publishing decisions and technological constraints. From this perspective, it is not the practice of reading that changes and shifts with the rise of digital media, but rather the support for meaning-making. The meaning of the text is altered when moving to digital texts: “one must state that forms produce meaning and that a text, stable in its letter, is invested with a new meaning and status when the mechanisms that make it available to interpretation change” (Chartier, 1994, p. 3). In sum, readers read more than the words (Topham, 2004).

Research into digital reading assesses the shift in meaning-making strategies by considering digital reading as a practice entirely different from paper-based reading. Screen-based reading
seems more volatile, and is characterized by “more time spent on browsing and scanning, keyword spotting, one–time reading, non–linear reading, and reading more selectively, while less time is spent on in–depth reading, and concentrated reading” (Liu, 2005, p. 702). According to Mangen, this bouncing around is due to the minimal mental energy required to click the mouse and rekindle our attention, which is easier than “to try to resist distractions by attempting to keep on structuring consciousness from within, and thus continue reading” (Mangen as cited by Hillesund, 2010, np). Bouncing from one site to the next, skimming only the first few paragraphs and bouncing out perhaps to never come back is typical of the horizontal information seeking of academic users, as exposed by the different CIBER studies (Nicholas, Rowlands, Clark, Huntington, Jamali & Ollé, 2008; Nicholas, et al., 2008; Rowlands, et al., 2008). Navigating becomes a fundamental aspect of accessing information online, instead of a secondary activity supporting the main reading task. Users tend to spend as much time finding their way around content as they spend on viewing their findings (Rowlands et al., 2008). This is indicative of a power browsing behavior: users move from one item to the next generally very rapidly, yearning for “quick wins” (Dillon, 2010). While users are often taken off course by power browsing, this behavior also leads to serendipitous findings, including interesting sites that could only be reached through such behavior (Hillesund, 2010).

Rowlands et al. (2008) claim that the short time spent viewing webpages or online article indicates that new forms of reading are emerging. Perhaps more intriguing is the “squirreling” behavior uncovered by Rowlands et al., and the numerous CIBER studies (Nicholas, Rowlands, Clark, Huntington, Jamali & Ollé, 2008). Academic users exhibit strong consumer instincts and “squirrel away content in the form of downloads, especially when there are free offers” (Rowlands et al., 2008, p. 295). While the extent to which these downloads are read is currently under-researched, this squirreling behavior may point to the importance of the materiality of the digital interface and the vast differences that exist between browsers, PDF viewers and other reading appliances.

We were unable to find in-depth research investigating the implications of bouncing, power browsing and squirreling on annotation and note-taking practices. The most common and general finding in this area of research mostly indicates that users who need to read in-depth will tend to print out their document and annotate on paper (Tenopir, King, Edwards & Wu, 2009; Liu, 2005; Shaikh, 2004). Annotating digital documents seems to require much more resources and skills on
the part of the user (McKnight 1997, as cited by Liu, 2005), and appears to distract the reader from the main reading task. Printing out documents for in depth reading is indicative of the fact that paper still dominates academic reading (Liu, 2005; Tenopir, King, Edwards & Wu, 2009). However, recent studies reveal that sustained screen reading may be increasing and that this may be due to the adaptation and adaptability of current academic users (Nicholas, Rowlands, Clark, Huntington, Jamali, & Ollé, 2008).

Dividing scholarly tasks according to the affordances of the medium (e.g., browsing and searching on the computer, in-depth reading and annotating on paper) is reminiscent of Sellen and Harper’s studies of the use of paper and computer in various organizations (1997, 2002). The authors observe how computer systems are normally preferred for making and remaking documents, storing, accessing, retrieving and distributing documents. Conversely, paper tends to facilitate creative tasks (e.g., editing, commenting and collaborating) and tasks that require sustained attention and concentration “such as reading, in which annotation, quick navigation and spatial layout of documents allow readers to deepen their understanding and to create a plan for their own writing.” (Hillesund, 2010, np)

2.2.2.3 Ways of Writing

Ultimately, relevance for writing guides the academic user in all aspects of information seeking and use. Research questions and personal motivations form the basis of the subsequent interpretation and meaning-making tasks and ultimately guide the selection process when highlighting, underlining or writing marginalia (Hillesund, 2010). Annotation and note-taking is intertwined with writing in multiple ways. Not only are annotations and notes traditionally used during the writing activity, but also specific types of annotations may be created when drafting and writing. This section investigates the interconnectedness of writing and annotating.

While the writing process is idiosyncratic, it often features the organization and refinement of notes and annotations into “a linear and/or hierarchical order appropriate for the type of document being composed” (Case, 1991, p. 78). An important step in the writing process is the creation of an outline. Often this emerges directly from annotations and generally quite early in the scholarly process: the outline guides subsequent note-taking, and vice versa. Donald Owen Case (1991) highlights the back and forth nature of scholarship, where scholars often juggle and alternate between multiple types of documents at once: source material, notes and drafts.
Writing has been described as the “complex interplay between the knowledge in the writer’s head and the physical and informational characteristics of the external source documents and compositions” (O’Hara, Taylor, Newman & Sellen, 2002, p. 280). This understanding of the writing process departs from the predominant psychological view of the phenomena, which does not generally address the influential role of materiality and technology. In their study of writing from multiple documents, O’Hara, Taylor, Newman & Sellen (2002) reiterate the importance of materiality and technology in fostering the back and forth nature of the writing practice. These movements of attention occur when comparing and contrasting information, copying, paraphrasing, referencing and integrating ideas and are facilitated differently according to the affordances of the sources documents. In order to enable this process, the user designs and arranges her workspace to have all necessary documents within reach, and adjusts this space according to the structure of her composition (p. 284).

The authors dedicate a large section to annotations created and used while writing from multiple documents. They reveal that the creation of annotations on source documents also occur in the writing activity, particularly on printouts of digital documents and photocopies. Annotation and note-taking might be essential in the writing stage: the cognitive activity of writing, which entails a complex interplay between the knowledge in the writer’s head, the source material and draft documents, may only be manageable by breaking it into discrete sub-tasks such as annotating, note-taking and outlining. This externalization of knowledge prevents crucial information from being lost.

Writing from multiple documents requires linking and navigating between different sources and documents. These links are often explicit: writers tend to make informal and incomplete remarks on documents linking to other sources (see section 2.1.2 Annotation as an Associative Device). The authors take a distributed cognition approach to the types of links and notes created in such an instance: navigational links (such as “see notes in the red book”) are a fitting example of how meaning is distributed across internal (what and where is the red book?) and external representation (the navigational link), and reconstructed when necessary (p. 290).

Finally, they briefly mention how users seem to be concurrently using paper and digital versions of the same document when writing, but for different purposes. Their results indicate that this dual-medium representation may be both helpful and inconvenient: different mediums may
support different types of search and thus may help structure the writing process. However, this increases the volume of documents to handle and perhaps also multiplies the possible location and types of annotation on a single article.

2.2.3 Intellectual, Social and Institutional Implications of Annotation

Recent scholarship on the history of the book investigates the social and institutional implications of annotation. These studies seek to understand the tensions between the reader, the book and the wider context of use. Annotation and note-taking are crucial clues for these studies as they shed light on the material practices and cultural expectations of particular readers. They furthermore reveal how specific modes of thought and argument, shaped by one’s intellectual milieu, predominated at a certain point in time (Blair, 2004b). This section addresses how the intellectual, social and institutional spheres influence annotation and note-taking strategies.

2.2.3.1 Annotations and Interpretive Communities

Some authors agree that, while annotation and note-taking practices may be idiosyncratic and overwhelmingly individual, they also follow consistent patterns across a certain genre or population (Blair, 2004a, 2004b; Daston, 2004). The importance of considering the reader as part of a larger reader community (or interpretive community) is the central tenet for many historians’ arguments, and forms the basis of the reading response theory. Annotation can therefore be considered as a material meaning-making strategy emerging from specific interpretive communities and social areas (Chartier, 1994).

Chartier (1994) suggests that there is no fixity of meaning; the significations constructed between a proposal (the text) and a reception (the reader) are always plural and mobile. However, books are produced within a specific order, implying rules, conventions and hierarchies. Chartier suggests that they nonetheless defy all of these. Books are deciphered on the basis of mental and affective schemes constituting the culture of communities that receive the work (p. x). Reading, as well as annotating and note-taking, is “shaped by cultural norms - transmitted more or less explicitly through education and imitation - and that can be differentiated according to cultural context” (Blair, 2004a, p. 421). Hence, the independence and freedom of interpretation is not only limited by the codes and conventions from the reading community, but also by the materiality of the texts (Chartier, 1994).
Chartier posits that the relationship between the reader and the book is indeed a mediated one (Chartier, 1994; Slights, 2001). He thus departs from the widely accepted notion of “horizons of expectations” as advanced by Jauss “where each texts a reader encounters automatically evokes previous texts and their expectations and rules” (as described in Colclough, 2007). This encounter between the signal emitted by the text and the horizon of expectations of the public is pure and unmediated by materiality or technology. Chartier (1989) and Cavallo & Chartier (1999) reject this notion when addressing the changing materiality of the book throughout the history of reading in the West. Reading, here, is a practice that is realized in acts, places and habits. Meaning does not only emerge from the readable space, but also from the form through which texts are received: texts shift in meaning with every change of material support.

The concepts of social areas (Roger Chartier), reading communities (Ann Blair, Stephen Colclough) and interpretative communities (Stanley Fish) are crucial to our understanding of annotation as a social practice influenced by communities and their prevalent modes of thoughts. Daston’s concept of “on-paper communities” addresses this explicitly. She describes readers perpetuating social codes and conventions by creating and using annotations for specific meaning-making purposes and to demonstrate their membership to the community. Daston speaks of the necessity of the note-taking process to not only retrace the arguments advanced by authors, but also, and speaking specifically of the habits of readers annotating Descartes, “to participate in his radical doubt and gradual restoration of belief. The apparatus of annotations, citations, and footnotes enmeshed both author and readers in a web of other works, with their respective authors and readers” (Daston, 2004, p. 447).

Interpretative communities are at the core of Grafton (1991) and Jardine & Grafton’s (1990) interpretation of Gabriel Harvey’s Livy. Their study unearths a new type of reader (the reader as facilitator) and reveals the importance of the centrifugal mode of reading. By studying his heavily annotated Livy, Jardine and Grafton revealed that Gabriel Harvey (c. 1545 – 1630), notable scholar whose reputation suffered from various quarrels, acted as a facilitator for noble men and politicians, “easing the difficult negotiations between modern needs and ancient texts” (Jardine & Grafton, 1990, p. 35). The annotations remaining in his book demonstrate that Harvey read his Livy multiple times and for very different purposes, but always as a “trigger for action” (p. 40).
His annotations show that his readings were connected to other texts, and that he even perhaps read (often alongside his colleague/employer) with “the appropriate books open on the table before them” (44). From the amount of works and books referenced, Grafton and Jardine posit that Harvey might have read using a book wheel (see figure 7). This demonstrates how Harvey was part of a wider community of readers, as the book wheel “belongs to Harvey’s cultural moment, in which collation and parallel citation were an essential, constructive part of a particular kind of reading; it allowed the imbedding of text in context, after the fashion that Harvey and (we would argue) many of his professional academic contemporaries practiced” (p. 48). The potential presence of the book wheel in his reading practice suggests that Harvey might have been perceived (or wanted to be perceived) as a “skilled reader”, just like other individuals may be skilled woodworkers. In sum, the annotations found in Harvey’s Livy are indicative of a centrifugal mode of reading, where the reader pulls information from multiple texts. This centrifugal reading practice is supported by the appropriate tools such as the book wheel, which situate Harvey as part of a wider community of readers.

Figure 7: Bookwheels and Their Users
From left, clockwise: the bookwheel from "The Various and Ingenious Machines of Agostino Ramelli"; bookwheel from Grollier de Servière's "Recueil d'ouvrages curieux de mathématique et de mécanique"; Anthony Grafton and his bookwheel
The role of the interpretive community in our study is linked to the social and disciplinary norms affecting the student’s practices. We posit that students are members of very specific interpretive communities (perhaps defined by disciplinary norms and boundaries), and that these communities constrain the annotation practices of student. Following in the footsteps of Grafton and Jardine, we believe that different social or disciplinary norms are observable in the materiality of the practice, not only in the notes and annotations, but also in the tools, support materials and in the layout of the reader’s workspace.

2.2.3.2 Annotations and Institutions

As discussed previously, reading and meaning-making are framed by the codes and conventions of an interpretive community. However, external forces and other contexts also impose various constraints. In his essay entitled “Reading as Poaching”, Michel de Certeau (1984) addresses the relationships that shape and control one’s reading. He dramatically opposes the blackboxing of meaning and the existence of a legitimate or proper reading. He consequently denounces this fiction created by the “social institution that overdetermines [the reader’s] relation with the text” (p. 171). Many contemporary authors (Nichols, 1991; Mayali, 1991; Slights, 2001) echo his sentiments, including Ralph Hanna III (1991) who directly assesses the influence of institutions on annotation practices. Hannah states that “questions of annotation always come back to issues of communities and institutions, and consequently questions of power” (1991, p. 89). De Certeau reveals that the reader may interpret more independently and acquire more freedom if these social institutions subside, since the “creativity of the reader grows as the institution that controlled it declines” (p. 172). This section collates research on the influence of institutions on annotations practices, more precisely the effects of academic disciplines and publishing standards on the annotator.

Academic disciplines and the prevalent attitude towards scholarship shape the way scholars read and annotate (Blair, 2004; Colclough, 2007). An interesting example of this is the humanist practice of commonplacing. Blair and Colclough retrace the consequences of these scholarly notebooks for scholarship up to this day.
The commonplace, a type of notebook collating selected passages from multiple sources categorized under appropriate headings, was a common scholarship methodology for humanists of the early modern period (Blair, 2004). Recent studies consider this practice as a primary tool for the organization of knowledge (Colclough, 2007). While this note-taking method was a common pedagogical instrument for school children and university students (Cavallo & Chartier, 1999), it also mirrored specific ways of reading and entailed a precise type of scholarship.

Early modern humanist tended to read extensively, probably due to then recent rise in the availability of books (Cavallo & Chartier, 1999). However, this overabundance of books caused by the democratization of printing (which was deemed harmful and dangerous [Blair, 2004]) also influenced the methods for keeping track of one’s reading. Humanist scholars and authors of the 16th, 17th and 18th centuries developed multiple strategies to cope with this information overload. In her thorough investigation of these various coping strategies, Blair (2003) states that the “overabundance not only fostered the diffusion and development of various aids to learning or "reference genres" but also affected the way scholars worked, from reading and taking notes to composing books of their own” (p. 12).

By studying commonplaces, researchers are able to assess the potential usage of these notes as well as gauge the general consequences of this method for note-taking and scholarship practices. The commonplace books were designed for retrieval: the presence of general headings facilitate the task of finding and refinding key passages that one might want to use in a discussion or to compose their own books (Blair, 1992). However, the method of commonplacing also dangerously affected citation practices. Indeed, common practice would dictate to copy passages while leaving out the author’s name. The sources became invisible in the commonplace book: all entries are equal and coexist with each other, with old facts residing near newer information.

The use of the commonplaces slowly declined towards the end of the 17th century. This is indicative of a shift in the type of intellectual work and in the values held in common by scholars and erudites. The rise of modern science yielded new methods of note-taking, with new formats, contents, and guidelines (i.e., full attribution to the source author was then reinstated) (Colclough, 2007). Scholarly work was then focused on understanding the argument of the text and not simply concerned with gathering appropriate quotations that could be later used in conversation or reworked as a new text.
This commonplace practice however had multiple consequences for scholarship and the printing industry before its decline. Printers facilitated commonplacing by providing typographical marks pointing out passages worth transcribing in the commonplace. The printing industry thus institutionalized commonplacing: printers pre-determined how a text should be read (Colclough, 2007). These printer-supplied typographical marks and the availability of reading aids and indexes were thought to foster a more passive engagement with the text. The “clear and unambiguous” text supplied by printers therefore didn’t require the active engagement of readers (Saenger and Heinlein, as described in Sherman, 2008)

According to the historians Saenger and Heinlein, this is evidence of a shift towards a passive reading on the part of scholars: individuals do not need to intervene in the margins anymore. However, this perspective strips the reader from her agency. Multiple authors, from Adler to Sherman, distance themselves from this line of thought and point to an even more active engagement on the part of readers. Nonetheless, this example demonstrates the influence of the printing and publishing industries on annotation practices and hints at the dynamic established between institutions and readers.

A second example of the impact of institutions on annotation practices concern the manicule, these little hands drawn in the margin of the text. William Sherman (2008) posits that manicules were among the most common symbols produced for and by readers. Drawn manicules have also been institutionalized: they disappeared from widespread use in the course of the 18th or 19th century and were replaced by printed fists. This transition from drawn to printed hands implies a radical shift in the producer of the manicule, its intended audience, and its functions. Whereas drawn manicules were used to index, classify, and literally point out passages worth noting, printed manicules rather point out sanctioned readings (Sherman, 2008; Slights, 2001).

William Slights addresses the influence of different institutions on annotations in his book *Managing Readers: Printed Marginalia in English Renaissance Books* (2001). While his argument primarily concerns printed marginalia and authorial annotations, Slights hints at the necessary infrastructure behind the production of texts which eventually influences the meaning-making strategies of the reader:

“Behind every printed book lurks a well hidden crew of agents. Partly human, partly institutional, and partly mechanical, these agents are responsible for the
production of an elaborately constructed artifact that stands at a great distance from that elusive chimera called the author’s intention” (p. 77).

Slight even considers editors, commentators, translators and printers “second authors”, shaping the meaning and reading strategies of the reader. Slight also notes how printers and the printing industry influence the annotation and note-taking practices of readers through the technologies used to produce texts. As seen previously, Chartier calls for an examination of the consequences of the materiality of the book on the meaning of the text, specifying that “whether in manuscript or in print, books are objects whose forms, if they cannot impose the sense of the texts they bear, at least command the uses that can invest them and the appropriations to which they are susceptible” (1994, p. vii-ix).

Studies of early print traditions (Slight, 2001; Jackson, 2001; Genette, 1997) indicate that many factors influence the materiality and functions of the book. This may be especially true of elements of the book, specifically the paratext (these items such as titles, running headings, margins and title pages) that contextualize the book in a material way (Slight, 2001). Slight posits that manuscript traditions, printing-house conventions and individual decisions greatly influence the form factor of books and constrain the reading strategies of individuals.

Heather Jackson, in her lengthy exploration of the marginalia of Renaissance readers, states that our contemporary customs and expectations are based on the traditional conventions and format of the book, as established in the early days of print. The use and functions of a blank space in a book is determined by its location on the page and in the book. Different places in a book command different functions. She notes that spaces at the start of a book tend to be used for ownership marks, blank pages at the beginning of a chapter would normally be used to summarize the content and conversely blank spaces at the end of a chapter are reserved for a general assessment of the chapter or section (as opposed to responses to specific statements). In essence, “marginalia mirror the texts they supplement” (Jackson, 2001, p. 41).

2.2.4 Zooming Out: Synthesis of the Section

This section addresses annotation and note-taking as part of the larger context of the scholarly practice. First, we situate annotation as part of the research project lifecycle. Information behavior literature, both descriptive and model-based studies, seems to have a very narrow view
of annotation in the context of the research process, either mentioning it only in passing, or even failing to mention such activity. Hence, according to most of the information behavior literature surveyed (White, 1975; Case, 1991; Chu, 1999), annotating and note-taking cannot be considered as a research-phase. Therefore, it may be more productive to assess annotation as a primitive or a feature of scholarly work. Shifting our perspective allows us to reveal the adaptable qualities of annotation, a practice seemingly intertwined with or, at least occurring concurrently with other major practices such as reading or writing.

Annotation and note-taking practices tend to vary considerably throughout the research process, according to the main activity of the research (i.e. reading vs. writing) and furthermore tend to be performed differently depending on the support medium (e.g., paper or screen). Different types of reading also lead to different types of annotations and note-taking activities (i.e. a reflective reading tends to yield marginalia, whereas a receptive reading may only generate highlights and reader’s marks). The materiality of the text and the medium that support its materiality (e.g. paper vs. screen) also influence annotation and note-taking strategies. Reading from screen tends to be shallower: users bounce around and squirrel away documents for future reading.

Annotating digital documents requires much more resources and skills on the part of the user (McKnight 1997, as cited by Liu, 2005), and distracts the reader from the main reading task. However, recent studies reveal that sustained screen reading may be increasing and that this may be due to the adaptation and adaptability of current academic users (Nicholas, Rowlands, Clark, Huntington, Jamali, & Ollé, 2008). Therefore, we might expect to see a greater proportion of annotation and note-taking strategies being performed strictly on screen. Paper and screen may also have complementary functions As pointed out by Sellen and Harper (2002), some scholarly tasks may be better performed on paper than on the screen (i.e. creative tasks seem to be better facilitated by paper). Conversely, the screen may foster more structural tasks such as making and remaking documents, storing, accessing, retrieving and distributing documents.

While annotation and note-taking practices may be idiosyncratic and overwhelmingly individual, they also follow consistent patterns across a certain genre or population (Blair, 2004a, 2004b; Daston, 2004). Interpretive communities guide the meaning-making practices and generate different modes of engagement (both intellectual and material) for the reader. Interpretive communities in the academic social world may be linked to social and disciplinary norms guiding the annotation and note-taking practices of students. However, other external influences
may guide the student in her scholarly tasks. Institutions such as the publishing and printing industries may shape the materiality of the text, therefore constraining the possible modes of interactions between the reader and the book.
3 Empirical Study: Annotation and Note-Taking Practices of Humanities and Social Sciences Graduate Students

The following section describes and analyzes the findings from the empirical study undertaken for this research. The broad aim of this thesis is to assess how annotation practices, as situated intellectual and material processes, are shaped by the multiple social worlds of the academic community. This study is primarily concerned with the different facets of annotation and note-taking, as situated practices in the research process. We therefore attempt to fulfill our three overarching objectives as previously described in the introduction (1.2.1 Research Questions): analyzing annotation as an epistemic practice, situating annotations and note-taking as materialities of infrastructure and finally locating the practice in the wider digital textuality context.

This first section examines the theoretical framework that guided the researcher throughout this study and details the instruments and procedures involved in the data collection process. The techniques and tools used for data analysis are then considered, outlining their filiations to the overarching framework and discussing their relevance and limitations. Results and analysis are then presented concurrently, framed by the multi-perspectival framework. This chapter concludes by discussing the implications of annotation and note-taking in the academic context and relates the findings back to the three overarching objectives.

3.1 Multi-Perspectival Framework

The framework designed for this study borrows from multiple theories of practice. It specifically draws from a pragmatist conception of practice and borrows from a materially-based understanding of epistemic activities. This multi-perspectival framework, focused on epistemic practice, follows Palmer & Cragin’s call for a more thorough exploration of scholarly activities in information science: "The holistic and materialist practice approach is well suited to research aimed at understanding the diversity of resources and activities involved in the scholarly process, but it has been thinly applied in IS" (Palmer & Cragin, 2008, p. 170). We thus present a framework borrowing from different theoretical traditions allowing us to zoom in and out of practice. This section explains the framework and its filiations to theories of practice and justifies
this assemblage in light of the current trends in critical information studies. After a brief contextualization of the framework in the wider realm of theories of practice, we delineate our different theoretical lenses borrowed from various epistemological traditions. This discussion leads us to finally highlight the different theoretical tools that are used in our analysis of the annotation and note-taking practices of graduate students.

3.1.1 Background

As noticed by many contemporary authors (Schatzki, Simpson and Bernstein for instance), the word “practice” is somewhat of a vague and ambiguous term, a buzzword of sorts, whose defining boundaries are fluid enough to permit multiple interpretations and significations and “imprecise and open enough to allow people from different traditions to join without renouncing their worldviews” (Miettinen, Smara-Fredericks, Yanow, 2010, p. 1313). Despite this vagueness and the lack of agreement on the extent of the concept of practice, most authors and thinkers agree that practice, simply put, is what humans do when they go about their daily tasks, the “embodied, materially mediated arrays of human activity centrally organized around shared practical understanding” (Schatzki, 2001, p. 11). The role of the material in the context of practice should be emphasized. Theories of practice centered on the active role of the material have been increasingly popular in the last few decades. Anthropologists and sociologists such as Bruno Latour (Jim Johnson) and Michel Callon made a concerted effort to understand the influence and consequences of materiality in social life. The material (or non-humans) does not only mediate human relationships and activities. In fact, the material is as equally emergent as humans (Schatzki, 2001, p. 20). From this perspective, concepts normally reserved for humans (agency, intention, purpose, knowledge, and voice) are equally important for non-humans.

In his 1972 book *Praxis and Action*, Richard Bernstein first traces the development of theories of practice to the Aristotelian notion of *praxis*. Praxis was then understood as the “disciplines and activities predominant in man’s ethical and political life” (p. xiv), which necessitated practical knowledge in order to “do” or “live well”, as opposed to *theoria*, which entailed the more contemplative knowing and wisdom. A second important element of the Aristotelian definition, that to this day transpires in contemporary notions of practice, is that praxis manifests itself in the polis and cannot be dissociated from it: practice is a social phenomena. Bernstein identifies four modern philosophical and theoretical endeavors concerned with practice, which recuperates
and integrates diverse aspects of the Aristotelian praxis: the societal theory of practice developed by Marx; the existentialist approach of Kierkegaard and Sartre showcasing the need to address action in terms of existence and consciousness; the pragmatist perspective concerned with practical action, and finally the analytical philosophers of Wittgensteinian tradition, addressing human speech as a form of action (Bernstein, 1972; Simpson, 2010).

The influence of theories of practice in social theory is qualified by many as the “practice turn” in applied areas of knowledge (Schatzki, 2001). Schatzki observes three main trends in the field. These trends re-emerge in the work of Barbara Simpson as she attempts to close the gaps between them and addresses how practices unfold in daily activities. These trends, or facets in our work, form the structural support of our framework.

The first trend engages with social orders and practices. The emphasis on the creation and maintenance of social structure is often opposed to the “human capacity for agentic action” (Simpson, 2010, p. 1331). The durability of socially constructed identities as well as theories surrounding communities of practice address this dualism and propose that practices are a dynamic movement towards a certain equilibrium between the social and the individual. The second contemporary approach focuses on the micro-perspectives of human activities, namely the “embodied capacities such as know-how, skills, tacit understandings, and dispositions” (Schatzki, 2001, p. 16). The embodiment of meaning is central to this category and joins other dispositions also in a convergent dynamic. Finally, and perhaps more interestingly for our framework, is the work of post-humanist writers such as Haraway, Latour, Callon and Knorr Cetina. These theories “engage with the dynamic of process and the functionalities of objects in creative meaning-making” (Simpson, 2010, p. 1332). These post-humanist approaches are not uniform or homogeneous, hence the use of the phrase “posthumanist challenges” by Schatzki (2001). Despite divergent ideas and theories, Schatzki presents two global varieties (Schatzki’s term) of post-humanism intersecting with practice theory. The first, termed “objectivism”, highlights how human activities are interwoven with non-human entities. The second agenda places practices over individuals. It understands individuals as emerging from their integration into social practice, since “there come to be people, that is to say, humans with activities, minds, identities, and genders through this incorporation” (Schatzki, 2001, p. 20).
We must then clarify how our framework stands in relation to these three trends by first looking at how they are addressed by both Schatzki and Simpson. These understandings, as well as ours, are visually represented in figure 8 and figure 9 respectively. On the one hand, Schatzki describes these as three separate and parallel trends in contemporary practice theories. They are linked to the authors’ specific understandings and disciplinary traditions. On the other hand, Simpson observes three diverging categories of practice theories. She specifically highlights the gaps between these categories and reveals that we may be able to address these gaps by examining how practices unfold in everyday activities.

Conversely, we look at these three trends as three facets of the same problematic: the practice. We present our understanding in figure 9 by using a cube analogy. While each side is distinct and different, they come together to form a unified shape. Therefore, we acknowledge the need to pull from various disciplinary traditions to gain a better global understanding of the practice. Additionally, edges and corners play a crucial role in supporting our framework (and the cube). We consider these edges as necessary structural elements, an invisible glue binding the different facets together. Therefore, the three sides described above (social order, micro-perspectives and post-humanism) are linked by a common thread, a structural element that allows practices to fully stand on their own.
Schatzki (2001) observes three distinct and parallel trends in contemporary practice theory.

Simpson (2010) understands these as three diverging categories. The gaps may be explained by addressing ‘how’ practices unfold.

Figure 8: Visual Representation of the Different Understandings of Schatzki (2001) & Simpson (2010)

Our framework understands different facets as both distinct and complementary.

Figure 9: Visual Analogy for the Structure of our Framework

It should be noted that annotations and note-taking practices require a framework that allows us to talk about these various faces and sides both separately and concurrently. We therefore need to distance ourselves from Schatzki and Simpson in our understanding of these three trends: while
they are still individual facets, they converge to create a bigger picture. This visual analogy are further explored in the next section 3.1.2 Zooming In and Out of Practice, as we consider how a bricolage of complementary disciplinary traditions can foster a more global understanding of the concept of practice.

3.1.2 Zooming In and Out on Practice: a Multi-Perspectival Framework

Our framework follows the recent suggestion made by Nicolini (2010) that, in order to fully comprehend the daily doings of individuals, one must both zoom in and out on practice by using different theoretical lenses. However, before describing the workings of this methodology, we must first define and clarify our vision of practice. Our conception of practice, and consequently of the activities, tools and tasks involved, is indebted to Meadian Pragmatism as applied to practice, symbolic interactionism and its recent resurgence in the works of Leigh Star, Bowker and Clarke, and the objectual and epistemic aspect of practice as advanced by Knorr Cetina.

Our choice of disciplinary and methodological traditions is in line with Schatzki and Simpson’s observations and, indeed, with the contemporary trends in practice theory. However, and as discussed previously, instead of approaching practice through only one of these lenses, our framework understands the necessity of addressing practice using these three complementary facets at once.

3.1.2.1 Pragmatism and the Practice Turn

In her paper entitled “Pragmatism, Mead and the Practice Turn” (2010), Barbara Simpson outlines a renewed understanding of practice based on pragmatism as advanced by George Herbert Mead. Simpson reveals innovative implications of Meadian pragmatism for practice, “engaging with ‘how’ practice emerges in real-time rather than ‘what’ practices are in use” (Simpson, 2010, p. 1343). The definition of practice put forward in this perspective, and adopted by our framework, is one where practice is the conduct of transactional life and involves “temporally-unfolding, symbolically mediated interweaving of experience and action” (p. 1338). This definition builds in many assumptions, which are regarded as tools or lines of questioning useful for this first study.

The conduct of transactional life is at the center of the pragmatist view of practice. *Transaction* is described by Mead and his colleagues as a conversation of gestures taking place between
actors who are part of a relationally integrated whole. These actors call upon significant symbols, gestures that evoke the same response in both the gesturer and the responder, in order to co-construct meaning. Transactional engagements allow for actors to explore the ambiguities in meaning, and indeed take advantage of these differences to evolve and expand meaning, as well as gain new insight.

While significant symbols may guide the expectations of anticipated behaviors, individuals ultimately make their own decision concerning their actions. However, these choices are also guided (or at least socially moderated) by significant symbols. Social habits of conducts, the organized system of symbols reflecting the general attitudes of a group or community, are known as the generalized other. While this may seem like a vague concept without any concrete application, Mead points to certain key actors that represent the generalized other, such as the bank teller representing ‘The Bank’.

Transactional events are furthermore temporally situated. Mead stresses the importance of temporality, not specifically in terms of a chain of transactional moments, but rather in terms of a “continuous narration of unfolding social selves” (Simpson, 2010, p. 1336). He acknowledges the existence of objective and subjective events. Objective events are essential for structuring the flow time. Time is only experienced when its flow is interrupted by unforeseen events, imposing into the timeline and creating new possibilities.

A second important aspect of temporality in Mead’s view is that “both the past and future are in the actions of the present” (p. 1338). The past in this situation is the multiple social attitudes and significant symbols in a social setting, while the future is often anticipated and enacted. This often leads to the possibility of creative action, due to the work of the subjective ‘I’, the “active principle of forward movement that introduces divergence and novel possibilities into the processes of the self” (p. 1336), while at the same time being informed by the ‘me’, the organized set of habits of conducts and other’s attitudes.

Finally, Meadian pragmatism overcomes pervasive dualisms in contemporary theories of practice, such as the separation of the individual and social, as well as the separation of habitual and creative actions. The former is effectively erased by the presence of the social self, the self emerging from transactions, continuously spanning the individual/social divide. The latter is
addressed by the dynamic processes of action, both convergent and divergent, due to the roles of the ‘I’ and the ‘me’ in a transaction.

Hence, our framework benefits from including Meadian theories of pragmatism in several ways. Significant symbols and theories of temporality allow us to zoom in on the materiality and the content of annotations, while the concept of the generalized other is useful to assess the institutional influence on the annotation practice. Furthermore, transactional events can become units of analysis allowing for the relationships between individuals, communities and institutions to emerge. Transactional events are not solely between two individuals: they may also involve larger bodies. This perspective fosters a greater understanding of the integration of the individual in a social circle, since the social self emerges from these transactions.

3.1.2.2 Symbolic Interactionism and the Bigger Picture

Simpson dismisses symbolic interactionism (SI) as a watered down, less innovative application of the Meadian theories. According to her, these theories are not of useful assistance to any researchers studying practice. While we agree that symbolic interactionism might focus mainly on the interactions at a micro-scale (leaving much of the important Meadian theories behind) we however disagree with Simpson and her reductionist stance on SI. We distance ourselves from Simpson and advance that symbolic interactionism, and more precisely its contemporary resurgence, is necessary for researchers of practice in order to observe and analyze the bigger picture. Authors such as Clarke, Star, and Bowker demonstrate the wider ramifications of SI with their theories addressing sociality and social worlds, ecologies of practice, and infrastructures and standards.

Hence, while we understand Simpson’s concerns, we consider the inclusion of SI theories and authors in our framework necessary on two counts. First, the focus on symbolically-mediated interactions reinforces the Meadian discourse on transactionality: this is of crucial importance when considering that notes and annotations are themselves symbols, often created as part of internalized transactions. Second, the contemporary revival of symbolic interactionism addresses a larger context and focuses on aspects such as social worlds and infrastructures that were mainly hidden in Meadian theories.
In her book on the underpinnings of situational analysis (SA), Adele Clarke (2005) reveals the complementary nature of the early radical Meadian approach and the more recent theories of the Chicago School social ecologies influenced by Mead via symbolic interactionism. The contemporary approaches put forth in the last few decades by authors such as Star, Griesemer, Bowker and Clarke influence our understanding of symbolic interactionism. This section reviews the aspects of SI that are used in our analysis in order to answer our research questions.

Symbolic interactionism (and especially situational analysis) provides us with the concepts and tools necessary to explore social situations as sets of actions, actors, and relationships loosely bounded by a common goal and framed by those in it and the researchers. We use the concept of the social world, blurring the SI and SA traditions. Social worlds give prominence to “meaning-making social groups” (Clarke, 2005, p. 109) and people doing things together, to paraphrase Becker (1986). These concepts intimately relate to Mead’s key notion of perspective and commitment: all actors have their own perspectives and habits vis-à-vis a specific situation or transaction. Social worlds and arenas are conceptual frames where the meso or organizational level of social life emerges and give voice to collective actors such as the generalized others, their work and their discourses. Social worlds may be constituted by generalized others, individuals (that may also be part of other social worlds), implicated actors (actors that are silenced or only discursively present), and non-human actors.

We note that our use of social worlds clearly recognizes the importance of non-human actors. Therefore, we side with Bruno Latour (Johnson, 1988) and his fundamental work on the sociality of non-human actors. Latour explores a type of reciprocal shaping between things, machines and individuals; a reciprocal shaping that is not unlike the relationship between culture and individuals. Hence, a clear consequence of the postructural analytical lenses of the last few decades is to place theoretical importance on things and materialities. While non-humans are attributed some type of agency, it is always set within the boundaries of the social agency of the individual, constructed by humans in specific situations (Clarke, 2005, p. 61). Pragmatist and interactionist discourses recognize the importance of objects and reveal the continuity between individuals and objects and showcase the experience of the self being bounded by physical things.
Zooming out from the materiality of the non-human, the notion of ecology frames the overall relationships between actors (humans and non-humans alike) and fosters a global understanding of the annotation and note-taking practice. The concept of ecology is often deployed in the social sciences. This may be attributed to the sociological tradition of the Chicago School and its use of ecological concepts mapping social life. The ecologies developed at the University of Chicago in the early to mid 20th century were literally mapped on the geographical notion of space and apprehended social data in terms of territory and geography (Zorbaugh, 1929; Park, 1952; Strauss, 1961). Their approach nonetheless set the agenda for future researchers studying practices and human life. Indeed, this geographical approach is often discernible in contemporary social studies examining the tension between the big picture and the local particularities.

Hence, we follow in the footsteps of the Chicago School tradition and of its contemporary applications. Our framework acknowledges the crucial work of Star & Ruhleder (1996), Star & Griesemer (1989) and Bowker & Star (1999) on social ecologies, boundary objects and infrastructures. We also note the productive application of ecological concepts to information study by Nardi and O’Day (1999) and the development of the information ecology metaphor in their book Information Ecologies: Using technologies with heart. While we can clearly trace their filiations to Chicago ecologies, these authors also represent a contemporary analytical shift in the understanding of ecologies and social life.

Through Leigh Star’s body of work, we come to understand the importance of infrastructure in supporting human organization. Our current understanding of infrastructure as “a fundamentally relational concept, becoming real in relation to organized practices” (Star, 2002, p. 116, paraphrasing Star & Ruhleder, 1996) refines and supplements the traditional definition of infrastructure as something in the background, invisible, “something that other things ‘run on,’ things that are substrate to events and movements” (Star, 2002, p. 116). By foregrounding the backstage elements of practice, or doing an “infrastructural inversion” (Bowker, 1994), Star and Ruhleder (1996) examine the properties of infrastructure as “that which is embedded; transparent; having reach or scope; is learned as part of membership; has links with conventions

---

3 See Clarke (2005, 40-45) for a summary of this movement, as well as fascinating visual examples of the progression from geographically-based ecologies to more social & institutional ecologies.
of practice; embodies standards; is built on an installed base (and its inertia); becomes visible upon breakdown” (Star & Ruhleder, as discussed in Star, 2002, p. 117).

Star and her colleague Griesemer assess the notion of institutional ecologies and examine this concept as a constellation of different social worlds that intersects at various points of passage where translation occurs and allows information, actor and objects to move from one world to another. Their efforts are based on the previous work of Everett Hughes who described and defined institutional ecologies, and imposed ecological analysis as an important methodology for the social sciences. Ecological analysis, according to Hughes’ approach, does not presuppose the primacy of a single viewpoint. The unit of analysis tends to be the enterprise or the situation altogether, not the individual.

The notion of shared conventions is closely linked to ecologies and social worlds. These conventions are embodied in boundary objects that inhabit several intersecting social worlds (Becker, 1982; Star & Griesemer, 1989, p. 393), and meet the informational requirements of both worlds at once. This notion is fundamentally present in Howard Becker’s body of work. Art Worlds (1982) attempts to define art collectivities and explicitly stresses the importance of shared conventions between different collective activities, especially those involving the artists, their peers, the audience and the support personnel. It is indeed the global system comprised of shared conventions and collective activities that creates the artwork, not the sole individual identified as the artist.

Recent literature demonstrates how these sociological concepts can be applied to other academic disciplines. For instance, Nardi and O’Day focus on the information ecology. They describe this variation as “a complex system of parts and relationships. It exhibits diversity and experiences continual evolution. Different parts of an ecology coevolve, changing together according to the relationships in the system” (Nardi & O’Day, 1999, p. 50). Multiple types of actors exist in this ecology, namely keystone species, actors necessary for the survival of the ecology, and mediators, whose work on the fringes help bridging across boundaries and translate across disciplines.

Hence, symbolic interactionism and its contemporary counterparts allow us to highlight the relationships created between different actors (humans, non-humans, communities or institutions) and to zoom out of the practice. Using this theoretical lens, we can also observe how
these relationships are modified and rearranged when ecologies and social worlds (that reciprocally shape and support these actors) are also modified and transformed.

3.1.2.3 Objectual and Epistemic Practice

The final trend influencing our framework stems from the work of Karin Knorr Cetina (1981, 2001) on the role of objects in knowledge-based practice. Knorr Cetina examines the creative and constructive qualities of epistemic practice that emerge when the individual confronts non-routine problems. The role of epistemic objects is crucial in this type of practice. Knorr Cetina defines these objects “by their lack of completeness of being” (2001, p. 185). This lack appears to be essential for scientific and knowledge-based practices as it provokes further questioning for scientists and researchers, which allows them to move forward in their work. Knorr Cetina (2001) therefore builds on the notion of epistemic things as described by Rheinberger (1997): those scientific objects at the center of the research process that are continually materially defined. Epistemic objects are far from being everyday objects, in the sense that they are continually prompting questioning because of their lack of completeness: they are not blackboxed like their everyday counterparts. Epistemic objects are therefore not only defined by what they are, but also by what they are not, by what they lack.

Epistemic objects exist simultaneously in a variety of forms – these multiple instantiations all partially complete each other, and stand in relation to the whole. They are not preparatory materials for the real thing, rather “it is the real thing itself that has the changing ontology which the partial objects unfold” (Knorr Cetina, 2001, p. 191). This dynamic, unfolding character parallels the “structure of wanting” characterizing the self (p. 194). Objects suggest which ways to look further “through the signs they give off of what they still lack” (p. 194). This lack fosters a lateral branching out of the epistemic practice as the “structure of wanting implies a continually renewed interest in knowing that appears never to be fulfilled by final knowledge” (p. 196), and implies the inevitability of divergent interests and conflictual breaks and forks.

3.1.2.4 Zooming In and Out of Practice

These previous concepts, lines of questioning and theoretical lenses allow us to both zoom in and out on practice in order to bring to the fore different traits and characteristics that would have been otherwise invisible. This, above all, is an attempt to complexify, or rather reveal the
messiness of everyday practice, situating our framework against all types of reductionism, without losing sight of our research goals. Our cube analogy still holds in the face of the different theoretical lenses and academic traditions previously discussed. Zooming in on practice entails, according to Nicolini (2010), looking at the sayings and doings of individuals, understanding the active role of material elements and infrastructure, the local methods and strategies of accomplishment, the body choreography, the practical concerns and assessing the sense and object of the practice. As seen in figure 10, zooming in entails looking at the individual sides of the cube, and examining how these are intimately connected to other sides, via edges and corners.

![Zooming in on different aspects: sides, corners, and edges.](image1)

![Zooming out by focusing on the overall shape, structure, and context.](image2)

**Figure 10: Zooming In (left) and Out (right) on Practice**

Conversely, zooming out involves tracing the association between practices, the reciprocal implications of practice (when a practice becomes a resource for another), the actors becoming mediators, the local and global effects of the practice and the effects of the global on the local. Careful consideration must be given to how methods, habits and tasks are learned, how one goes from novice to practitioner, and how certain social processes foster this learning. We must also turn our attention to other people who perpetuate and make the practice durable, to the community sustaining the practice “as long as we agree that it is the practice that generates the social relations which emerge around it and not vice versa” (Nicolini, 2010, p. 1406). This is visually translated in figure 10 (right side) using our cube analogy. Zooming out thus involves looking at the overall shape of the cube, the specific arrangement of the sides, the invisible structure supporting the shape and finally, looking outside the cube to see its context and other cube-like shapes supporting it.
This framework involves an intensive back and forth between zoomed in and zoomed out states and therefore seems appropriate for the practices of annotation and note-taking. Our literature review points to the multiple angles and sides of annotation and note-taking, each addressed from different perspectives and scholarly traditions. While each of these perspectives unearths important findings, we notice a global lack of comprehension of the practice and a failure to explicate how the different levels of granularity of this practice are connected. Indeed, studies addressing the materiality of annotations do not consider the repercussions of this materiality on larger processes such as reading and writing or again on institutional representations of the individual as a student, a researcher or an assistant. Conversely, studies that examine larger processes involving annotation often fail to explicate the connections to more physical elements of the practice such as the workspace, the tools, and the body.

Hence, this zooming in and out is necessary for our study of annotative practices of graduate students. Our three objectives (analyzing annotation as an epistemic practice, as a materiality of infrastructure and situating the practice in the wider digital textuality context) indeed require different levels of granularity in our analysis, as seen in figure 11. We note that these objectives do not reside only at a single level of granularity. While they may be primarily concerned with a certain level of granularity (e.g. the materials and tools of annotation, a very low-level concern), these objectives require different theoretical lenses and levels of granularity on order to be fully addressed.
3.2 Methods

Following in the footsteps of previous studies of humanities and social sciences scholars, this study attempts to grasp the concept of annotation in the scholarly context through semi-structured participant interviews supplemented by document collection. The aim of this data collection is to gather significant information on the experiences of graduate students in the context of their daily activities by focusing on the role and importance of the practices, activities and materials involved in the research process. This type of qualitative data gathering allows the researcher “to identify elements of the work process, discern variations in information practice and use, and determine the relative value of materials and activities, largely from the respondents' point of view” (Palmer & Neumann, 2002, p. 91). The focus on the inner experiences of the participants lets the researcher assess how meanings are constructed through and in culture, and ultimately leads her “to discover rather than test variables” (Corbin & Strauss, 2008, p. 12).

3.2.1 Sampling Procedure and Participants

Participants for this study were selected according to multiple criteria. This purposive sample allowed the researcher to actively seek individuals who would provide rich accounts of their
scholarly practice. Purposive sampling was necessary for this research, since the topic and scope of study called for very specific qualifications in participants. Hence, we wanted to select “information-rich cases” (Patton, 2002, p. 230): cases or participants that allowed us to “learn a great deal about issues of central importance to the purpose of inquiry, thus the term purposeful sampling” (Patton, 2002, p. 230). We should note that this type of sampling prevents us from generalizing our findings outside of our population. However, while this limitation may have serious consequences, we believe that this sampling procedure yielded more in-depth findings and insights that may not have been detected if we used a more standardized, probabilistic sampling.

To this effect, all selected participants were graduate students recruited from social sciences or humanities departments. The choice of students rather than well-established scholars can be justified in two ways. First, graduate students tend to be more reflexive about their practices, often due to the shift to a higher level of scholarly activity (i.e., from undergraduate to master’s or master’s to doctorate). Second, this reflexivity leads to an adaptation of methods and techniques deployed in their daily activities as the students are confronted with new intellectual situations. The selected participants were all in the midst of a major individual research project, either completing a master’s thesis, submitting a major grant proposal, completing an independent study or finalizing a dissertation proposal. This second criterion ensured that participants created their own research process as compared to following required procedures for coursework or assignments. Furthermore, focusing on a single research project allowed the researcher to assess representative events and incidents that were ultimately linked by the common thread of the project at hand.

Seven participants were thus recruited from a large Canadian university: four were completing their Master’s degree and three were working towards doctoral candidacy. The researcher first contacted the potential participants via e-mail to assess their intention to participate (see Appendix A for recruitment letter). The participants were selected from departments and Faculties who theoretically position themselves as overlapping humanities and social sciences. The research interests of these students as a group thus spanned multiple fields, methods and perspectives. However, while their interests diverged, all were studying some aspects of recent technologies, either in order to apply their findings to the development of new technological solutions, or to situate the rise of different technologies within different historical or social
contexts. Finally, the sample consisted of 3 female and 4 male participants. Table 6 summarizes the main characteristics of our participants. It should be noted that to respect our participant’s privacy, pseudonyms are used throughout the study.

### Table 6: Brief Description of the Participants

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Gender</th>
<th>Age</th>
<th>Current Educational Level</th>
<th>Current Research Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meghan</td>
<td>Female</td>
<td>25-30</td>
<td>Doctoral Student</td>
<td>Writing SSRHC proposal; Preliminary Research for Comprehensive Exam; Preparing 3-year Research Project with Supervisor</td>
</tr>
<tr>
<td>Chris</td>
<td>Male</td>
<td>25-30</td>
<td>Master’s Student</td>
<td>Finishing up coursework; Writing Research Agenda</td>
</tr>
<tr>
<td>Alex</td>
<td>Male</td>
<td>30-35</td>
<td>Master’s Student</td>
<td>Completing Master’s Thesis Research; Writing Master’s Thesis</td>
</tr>
<tr>
<td>Mark</td>
<td>Male</td>
<td>25-30</td>
<td>Doctoral Student</td>
<td>Writing SSRHC proposal; Preliminary Research for Comprehensive Exam</td>
</tr>
<tr>
<td>Maxine</td>
<td>Female</td>
<td>30-35</td>
<td>Doctoral Student</td>
<td>Research for Comprehensive Exam</td>
</tr>
<tr>
<td>Sophie</td>
<td>Female</td>
<td>25-30</td>
<td>Master’s Student</td>
<td>Completing Master’s Thesis Research; Writing Master’s Thesis</td>
</tr>
<tr>
<td>Jacob</td>
<td>Male</td>
<td>25-30</td>
<td>Master’s Student</td>
<td>Completing Master’s Thesis Research; Writing Master’s Thesis</td>
</tr>
</tbody>
</table>

### 3.2.2 Research Instruments: Interviews and Document Collection

#### 3.2.2.1 Interviews

Individual interviews, following an open-ended and semi-structured format, were conducted with the participants in their natural environment. This type of “conversation with a purpose” (Kahn & Cannell, 1957, p. 149) allowed the researcher considerable freedom to probe beyond the established interview guide and down a topic of interest if a participant was particularly responsive. This method also let the researcher counter a lack of response by adapting and rephrasing questions (Herman-Kinney & Verschaeve, 2003, p. 231). The questions and topics, asked in a general and nondirective manner, prompted individuals to retell events and incidents in their own terms. Thus, the interviewer walked a fine line between eliciting the appropriate data for the research by “springing” participants and overspecifying the substance of the talk (McCracken, 1988, p. 34).
The lines of discussion and questions used in the interviews stemmed from preliminary research and from the ongoing data collection. This type of theoretical sampling (i.e., sampling for events, incidents, concepts and happenings) is in line with grounded theory, which interrelates data collection and data analysis (Corbin & Strauss, 1990, p. 8). Thus, data was analyzed concurrently with its collection; emerging themes and categories served as a basis for future participant interviews. Even if the interview guide was modified considerably according to emergent themes and to the personal characteristics of the participants, a certain structure remained. An example of the general interview guide used during these sessions is available in Appendix B.

Interviews normally opened with “grand-tour” questions (McCracken, 1988, p. 35) to elicit the participant’s academic background and field of interest. The line of questioning then moved towards the participant’s individual research project. The interviewer used floating prompts in order to encourage participants to expand on specific topics as the interviewee brought them up naturally. Alternatively, planned prompts, such as asking the participant to recall and describe specific events, were used to pinpoint certain aspects of the scholarly process. Finally, the interviewer turned to documents used in the research process of participants to elicit specific information on their material practices. The participants were asked to describe these documents and recall when and how they were created or modified. These documents, when possible, were also collected to supplement the data collected during the interview. These documents are considered as artifacts of the research process, enabling the researcher to also obtain the language (both from a linguistic standpoint, as well as in terms of the marks used to annotate) used by the participant (Creswell, 1998, p. 187).

3.2.2.2 Document Collection

Two types of documents were collected: physical evidences of textual practices and communications between different actors.

Textual practices have physical manifestations (e.g., notes and annotations) and are also present in the written communications between actors of the same social world (via e-mail, forum postings, comments on blog postings). It was therefore necessary to collect such evidence to analyze the content of these manifestations and to derive a classification scheme of the different textual practices. For instance, annotations may be personal reminders to further explore a topic
or a one-way dialogue with the author: the distinction was crucial for classification of the textual practices.

Evidence of communication between and within different social worlds was also of interest to this research. Such documents may be e-mails, corporate communications such as brochures and websites or instruction manuals for devices. Analyzing these documents helped the researcher gain further understanding of the enrollment process between different social worlds; the language used revealed how interests of one social world become aligned with another.

3.2.3 Procedures and Data Recording

The interviews were conducted in the normal work setting of the participant in order to gain greater understanding of the physical consequences of the annotation practices (e.g. looking at book shelves and classification schemes, bookmarking methods used, layout of the work or lounge areas, relation of digital devices with physical books, etc). The interview sessions averaged forty-five minutes and were audio recorded upon consent of the participants. The audiotape of the interview was then transcribed for analysis. Because the conversation occurred in the participant’s work setting during the regular course of her activities, the use of a video camera could have hindered the flow of a session and may have acted as a barrier between the participant and the interviewer. Since the natural setting and body language of the participant was of high importance to this research, the researcher kept an interview log, jotting down interesting elements of the setting and of the participant’s behavior.

Documents found in the participant’s setting were then collected in order to conduct a full textual analysis. These documents, in printed or digital format, were either scanned or copied onto the researcher’s flash drive. In some instances, for digital material that couldn’t be copied directly (e.g., a file organization system on a desktop computer), the participants were asked to provide screen captures of their desktop or applications. Electronic communications, such as e-mails, were forwarded to the researcher upon consent. The researcher then copied the content of the e-mail to a text file for future reference and deleted the original source to avoid the dissemination of any private information. Other types of electronic documents such as websites, forums and blog postings were downloaded and saved in a PDF format allowing the layout to remain consistent as well as adding a timestamp to every page converted to PDF. This is especially useful given the fluid nature of this type of material.
Audio recordings of interview sessions, as well as transcripts and any other digital information are kept solely in password-protected folders on the researcher’s personal computer, which is also password-protected. These recordings and transcripts will be stored for a maximum of two years and will then be destroyed, in accordance to the University of Toronto’s guidelines on data security. Access to any of the aforementioned data will be restricted only to the researcher and her supervisor.

3.2.4 Data Analysis

Data was analyzed using a strategy similar to grounded theory. Audio transcriptions and visual documents were analyzed first using an open coding technique. Open coding lets categories emerge (Babbie, 2001, p. 366) and is thus less directly related to the researcher’s preconceived social constructs. As coding progressed, a constant comparative method was used in order to saturate the categories by constantly referring the fresh data to emerging categories (Creswell, 1998, p. 57). Tams Analyzer (Weinstein, 2008), an open source qualitative analysis tool, proved to be useful for this first analytical step.

Once the open coding ended, the researcher engaged in axial coding, looking for “conditions, strategies, interactions and consequences” (Strauss, as cited in Berg, 1998, p. 240) and assembling the emerging categories in new ways. Here, evidence of recurring themes were gathered, data was searched for the various strategies and aspects of the annotation and note-taking practices, and emergent themes were compared to themes and constructs gleaned from the literature review. Additionally, the researcher noted the tone and vocabulary used during the interviews, adding to the general discourse analysis of this research. This deductive approach provided the foundation blocks for answering our research question.

3.2.5 Ethical Considerations

Since this research dealt directly with human subjects in their natural settings, the researcher took due heed to respect the needs and rights of the participants. Qualitative methods are often highly intrusive and the researcher may encounter sensitive information related to either the personal interests of participants or to their institutional activities. In order to reduce these risks and to protect the participants’ rights, the following steps were taken (adapted from Creswell, 1998, p. 202):
1. The researcher obtained full approval (delegated review) of the Social Science, Humanities & Education Ethics Review Board (ERB) at the University of Toronto before undertaking any fieldwork (Protocol Reference #24622, valid from November 20, 2009 to November 19, 2010). A copy of the ethics approval is available in Appendix C.

2. The research goals, as well as how the data will be used, were fully disclosed to the participants, both verbally and in writing.

3. The researcher obtained written consent from the participants agreeing to the research description before beginning any fieldwork (see Appendix D for letter of consent).

4. The participants were informed of all data collection techniques used (audio recordings, note-taking, observation procedure, photograph of artifacts and documents, document collection).

5. Original data, transcriptions, written interpretations and reports were made available to the participants.

6. Any identifying information was recorded separately from the research data to protect the anonymity of the participants and institutions involved.

7. Participants and institutions are referred to by using surrogate names throughout this report.

3.2.6 Limitations of the Methodology

While our methodology fosters a deeply focused understanding of the issues central to our study, it also impacts on our results due to its limitations. As discussed previously, several limitations stem from our sampling strategy. First, purposive sampling restrains the generalization of our findings to the selected group or type of participants. This focused approach, coupled with the characteristics of our research instruments, also prevent us from surveying a larger population. This lead to undercoverage: our small sampling of graduate students is by no means representative of the total population of graduate students. Selecting an appropriate sampling strategy is a series of trade-offs. In this case, our objectives and research questions called for a more in-depth understanding of epistemic practices. While a wider sample would have allowed more generalization, it would however have hampered a more finer-grained assessment of daily practices due to the time-consuming nature of our research instruments.

Our selected research instruments also have some inherent limitations. While they allow us to obtain more finer-grained findings than a survey or other type of quantitative instruments, they fall prey to the regular shortcomings of qualitative strategies. First, the use of self-reported data is problematic: the indirect information is filtered through the lens of the participants. However, we attempted to counter this limitation by supplementing interviews with document collection. The documents amassed contained traces of the practice, annotations and notes written during
the regular research activities of the student. Second, the presence of the researcher during the interviews may guide or bias the interviewee. In order to counter this negative effect, we constructed an open ended and informal interview guide that was adapted on the spot, according to the interests of the participants. Finally, interviews and document collection do not give us access to the performance of the practice itself. Perhaps shadowing or observing our participants throughout their daily activities, for an extended period of time, would have provided the perspective necessary to gain a greater understanding of the practice.

3.3 Results and Analysis

Our findings reveal multiple variations in the annotation and note-taking practices surveyed in the field. From the tools used and the organization of the notes, to the material preferences and the type of information encoded, the students truly develop their own idiosyncratic strategies for annotating material. However, despite these very personal and particular strategies, broader patterns of creation and use can be recognized. This section reports on the results from our interview sessions and document collection and sets the stage for further analysis. We first zoom in on the characteristics of the annotations and notes found in the field. We address their content, format and the tools used for creating and managing these notes. We then zoom out to expose their purposes and uses within the scholarly context, highlight the social and academic norms influencing the practice and finally hint at a potential micro-cycle of creation and use of these artifacts.

3.3.1 What do Scholarly Annotations and Notes Contain?

We approach these results with our general definition of annotation in mind, that of “any extra information about another piece of information, either attached to a document, or found on an external media.” These extra pieces of information differ greatly for our participants, both in terms of content and type. This research reveals multiple types annotations and notes. Whereas previous research (Marshall, 1997; 1998; 2009) has yielded a rich typology of annotation, our research offers a more focused typology centered on the types of information necessary to the completion of the academic research project.

We discuss and analyze several findings in this section. First, we examine the different types of annotation unearthed during our interviews and document collection. This typology expands the
findings of Marshall (1999), Ovsiannikov, Arbib & McNeill (1999), Agosti, Ferro, Frommholz & Thiel (2004), O’Hara (1996), and Blair (2004) by paying special attention to the role of meta annotations and notes (i.e., notes other than the ones containing research-specific content). We then elaborate on this typology by proposing a classification along a content-based spectrum.

Switching gears, we situate this typology within the context of the research project, observing the creation and use of various types of annotations and notes during different research activities. We thus expand on the typology of reading goals as posited by O’Hara (1996) and support the recent findings of the various CIBER studies regarding the rise of different information-seeking behaviors in an online environment (Nicholas, Rowlands, Clark, Huntington, Jamali & Ollé, 2008; Nicholas, et al., 2008; Rowlands, et al., 2008). We further distance ourselves from the previous literature on annotation by considering the creation and use of annotation during the course of activities other than reading.

Finally, we assess the function of annotation as linkage. We discuss this type of annotation separately due to its profound repercussions on the research project. While students are creating links between different documents by using explicit codes (e.g., “see page 23”, “cf”, etc), we also believe that the hypertext created during these events is crucial for moving the student towards analytical activities. Hence, this type of annotation performs various functions: linking between different documents, easing the recall task and alleviating the superfluous navigation to and from various documents, and finally bridging different research activities.

3.3.1.1 Common Types of Annotation in Research

Despite their idiosyncratic qualities, our participants’ annotation can be classified into broad categories. The following typology is based on the content (both textual and visual) as well as the formal aspect of the annotations and notes created by our participants. Figures 14 to 20 illustrate the most common types of annotation in our sample.

3.3.1.1.1 Selections
This type of annotation manifests itself frequently and appears in different formats. While all annotations are indeed a form of selection of the original source document, this type of annotation visually demarcate the selected passages and illustrate the scope of the annotation.
Two major types of selections exist: 1) sentences and phrases selected via highlights and underlines, as seen in figure 12; 2) copied and pasted passages, or quotes transcribed in an external document as illustrated in figure 13.

Selections, such as the ones created by underlining or highlighting text, are often used concurrently with a second method. In this instance, the highlight or underline is not necessarily considered as an annotation in itself. Rather, the use of this visual marking is only a part of the larger annotative act: the underline or highlight is simply the anchor element of the body-anchor-
marker anatomy as described by Marshall (2009). Figure 14 illustrates the use of highlights as part of a wider annotative act. In this example, Alex highlighted a bibliographic reference that appeared to be wrong. He included the correct information in a note situated close to the highlighted entry.

Figure 14: Example of a Highlight Used as Part of a Larger Annotative Act

While such strategies are present in the field, the use of standalone selection (i.e., the use of a highlight or underline by itself, without any labeling) is more prominent. Hence, the additional information appended to the original text is the marking itself: the highlights and underlines visually demarcate the important passage from the rest of the text. This demarcation conveys the importance of the selection. Quotes copied and pasted to other documents work similarly. The phrases detached from their original context and pasted into another document confer a certain importance of the chosen text. We note that the function of this type of annotation cannot be entirely defined here since highlights and other types of selection perform multiple functions according to the context of use. The functions of such annotation are further explored in section 3.3.4. Annotations and Note-Taking Lifecycle.

3.3.1.1.2 Emphasis

Emphasis is often conveyed through graphical means, whether with a cluster of asterisks, an arrow pointing to a paragraph, or the bold weight of a word in an already-highlighted phrase. We note that this technique is often used in addition to other types of annotation, as seen in figure 15 where a red arrow emphasize an already-highlighted passage.
3.3.1.1.3 Comments

Comments engage with the source material. They are either attached directly to a portion of the text (i.e., directly related to a paragraph or a section), or relate to the text in its entirety (such as general assessments of text often found at the beginning or end of the text). This is described in previous research as a “dialogue with the author” of the source material, and is the core interaction in active reading (Kopak & Chiang, 2008; Adler 1940a; 1940b).

Alongside the more explicit and fleshed out comments, this category also contains two other sub-types: questions and interjections. Questions generally target either the content, the author or the student’s own understanding of the text (e.g. “what does this mean?”). Conversely, interjections take the form of short written or symbolic notes made rapidly as a reaction to the text under study: an aggressive “NO!” in the margin or an exclamation point generally fulfill this task. We distinguish these sub-types from the more fleshed out comments in order to highlight the large spectrum of self-involvement and cognitive costs present in this category. While all written annotations require the student to break away from the main reading task, a long marginal annotation necessitates a rather prolonged departure from the text. This is a significantly different cognitive cost compared to an interjection that can be added almost simultaneously to reading.

These types of comments are the humanities/social sciences counterpart to the category “In situ way of working problems” described by Marshall in her study of physics textbooks (1999). Here, rather than having equations, variables and formulas in the margins, the student works through
the author’s argument by pointing out the weaknesses of the text, emphasizing the important ideas and writing down related thoughts that emerge in her head in order to comprehend the text. Figure 16 illustrates this category by showing Meghan’s reactions to different passages in a text.

3.3.1.1.4 Interpretive Notes

Conversely, the interpretive notes category identifies comments that do not necessarily relate to a specific source, but that rather expose the thought process of the individual. The notes in this category span a full spectrum of self-involvement, from the thoughts and ideas that seemingly appear out of the blue (such as the ideas for a theme or topic to explore, rapidly jotted down on a scrap of paper) to the more structured and organized thoughts looking at the ‘big picture’ (such as the notes necessary to the construction of an outline). Figure 17 shows the early stages of Alex’s outline for a chapter of his thesis. This example intertwines comments of wider scope
(headings and titles for different sections) and specific notes related to certain points that Alex needs to address in his final text.

**Figure 17: Example of Interpretive Notes in the Form of an Early Project Outline**

This category hints at a clear distinction between notes and annotations, two terms used interchangingly to this point. While ideas and thoughts can be written in the margins of a source document, the fact that they do not necessarily relate to the text or add any information to the source places them in the “notes” category. This is in line with our definition of *annotation* as an “explanation or comment added to a text” since this definition presupposes a certain relationship between the note and the text.

### 3.3.1.1.5 Explicit Reminders

The last category relates to a more *meta* type of information typically found in research projects. These notes often take the form of reminders, deadlines or to-do lists related to the project, such as the note in figure 18 assigning the task of finishing Chapter 2 by a certain date. Interestingly, students create these notes on a wide variety of material and supports. Hence, students do not only append meta notes to source material or create them in notes files (see “To-do” note in
Alex’s file, figure 19). Project-related notes also appear in calendars (such as figure 20), agendas, scrap papers, and sticky notes.

Figure 18: Example of a Project-Related Meta Note in a Calendar

These notes often relate to the overall project, the research activities and tasks, or the content and material of the project. From project-related to content-related, multiple sub-types run the gamut of meta notes, ranging from stylistic and grammar notes (e.g., proofreading marks), to notes concerning the overall project (e.g., “Submit Chapter 1 Tuesday!!”), activities and tasks (e.g., call numbers), and the content itself (e.g., “Delete second paragraph?”). See figure 19 for a visual representation of the gamut of the meta notes, including indicators for each of the sub-categories.

Figure 19: Diversity of Meta Notes on a Spectrum Ranging from Project to Content-Related
3.3.1.1.6 Notes on this Typology

This typology points to two larger categories of use in the context of the research project: research notes and meta notes. This overall categorization sets our study apart from previous research on annotation (Ovsiannikov, Arbib & McNeill, 1999; Agosti, Ferro, Frommholz & Thiel, 2004; O’Hara, 1996; Blair, 2004) in the active consideration of meta notes as a crucial element in the research process, alongside research notes. While Marshall (1999) considered incidental notes as part of the reading process of students, we highlight the fact that meta notes seem to gain importance in research that is beyond their incidental nature. Their constant presence throughout the research project suggests that they may act as an important support strategy for young researchers, helping them accomplish academic tasks and activities that are generally new to them. Perhaps, then, these types of notes are more in line with the role information scraps in the daily activities of knowledge workers as discussed by Bernstein, Van Kleek, Karger & Schraefel (2008).

We also noticed two sub-categories of research notes: reading notes and personal (or interpretive) notes. Figure 20 illustrates our new typology. While this typology is in line with previous research and indeed borrows similar categories, it also departs from these studies in several ways. First, this typology is based on the perceived higher-level functions of the notes and annotations as surveyed in the field and takes into consideration the context of creation. This perspective departs from studies that assess annotations in terms of higher order thinking elements (e.g. Ovsiannikov, Arbib & McNeill, 1999), or in purely functionalist terms (Agosti, Ferro, Frommholz & Thiel, 2004; O’Hara, 1996). Second, our typology is restricted to personal research projects. This allows us to distance ourselves from Marshall (1999) and her study of annotation in class and textbook contexts, and from Agosti, Ferro, Frommholz & Thiel (2004) and their research on annotation in collaborative setting. Finally, we recognize the fluidity of these labels and the porous boundaries of our categories: notes and annotations may have multiple intertwined purposes. For instance, comments may contain both reading-related material and interpretive material. We thus acknowledge that our typology may be more efficient if presented as a spectrum ranging from meta notes to interpretive content (see figure 21).
Selections in the form of an underline or a highlight, as well as emphasis markings are difficult to classify in this typology due to their graphical nature: without any textual content, one can only guess the ultimate purpose of the marking. Their visual qualities, however, allow for these markings to be layered on top of written content (the source material or other comments or annotations). In fact, they can only gain their meaning if appended directly to the material: the type of material to which they are added ultimately determines the general purpose of the marking. Hence, graphical markings are extremely versatile. They attract attention to specific portions of the text while concurrently being vague and open-ended enough about the specific
type of information they add to the original material. For this reason, we situate these markings outside of the meta or research categories (figure 20), since they may belong to, or reinforce either.

3.3.1.2 Unearthing Diversity: Types of Notes and Annotations Throughout the Research Project

While most of the participants have multiple annotating and note-taking strategies in common, this study also reveals idiosyncratic activities that, while they may be personal and certainly not generalizable, however hint at the pervasiveness of annotation and note-taking throughout the research process. The following sub-section looks at how different research-related activities foster specific types of notes and annotations and how, vice-versa, different types of notes and annotations may define these research activities. Further analysis of the research process can be found in section 3.3.5 How does the research process affect the nature of annotation and note-taking, where we take a global look at the research process and suggest that the creation and use of different types of annotations and notes allow the student to progress from one activity to the other. Appendix D summarizes the types of notes and their characteristics, according to the activity that fosters these specific strategies.

3.3.1.2.1 Ideation and Early Thought Development

Our study reveals that quick jottings are very frequent in the early days of the research project as the student attempts to establish the theme, scope and boundaries of the research. It is also at this stage that participants such as Maxine create visual overviews of their current knowledge and of their prospected research through mindmaps. Participants also identified the need to quickly copy and paste citations and bibliographic information to text files as they were researching background information on their topics. Mark further talked of quick jottings on whiteboards as a result of meeting with his advisor at the start of a new project. These quick notes are often transferred to another medium (if deemed important enough) in order to ensure their reliability and longevity. This study reveals that very few annotations are created at this stage. Rather, notes are the prevalent mode and are often very temporary, discarded as soon as their usefulness recedes. This ephemerality is linked to the type of information necessary for students at this stage of the research. Students require information to help them move on to the next stage of the
research normally comprised of reading primary and secondary materials and conducting research.

It is crucial to note that, conversely, some documents that started with a few words in the early days of the project can follow the student until the completion of the research and be regularly updated and used throughout the project. This is true for the concept maps created by Maxine, which slowly morphed into an outline:

>This is kinda weird. So, when I first start, I usually make a concept map and then I go from there. Because I have trouble organizing my thoughts, so the concept map is kind of a guide to determine which areas to really focus on. So from there, I’ll probably get my outline, and then start to write my outline and then key points and key areas.

The information contained in these documents are often re-worked and re-organized as the project progresses. These documents often bear the traces of the many steps leading to the completion of the project and visually represent the student’s research process. This type of document is very frequent among our participants. We further explore and analyze these documents (that we named intermediary documents) in section 3.3.4.3 Transfer.

3.3.1.2.2 Reading Primary and Secondary Source Materials

When prompted to discuss the different types of annotations made during their research, participants often referred to notes created while reading. These notes commonly parallel the ones described in the previous section 3.3.1.1 (most notably selections, emphasis, and comments). However, different types of reading require different types of annotation and note-taking strategies. The perceived importance of the source material (e.g., if it is a seminal article or not), as well as the academic provenance of the article (e.g., journal type, discipline) furthermore influences the depth and breadth of annotations. The following typology surveys the types of reading encountered in the field. These types of reading complement the typology of reading goals as developed by O’Hara (1996). Table 7 compares our typology of reading to O’Hara’s reading goals.
Table 7: Supplementing O'Hara's Typology of Reading Goals (1996)

<table>
<thead>
<tr>
<th>O'Hara's typology (1996)</th>
<th>Complementary types from our research</th>
</tr>
</thead>
<tbody>
<tr>
<td>How a text is read</td>
<td>Why a text is read</td>
</tr>
<tr>
<td>Receptive reading;</td>
<td>To learn;</td>
</tr>
<tr>
<td>Reflective reading;</td>
<td>To self inform;</td>
</tr>
<tr>
<td>Skim reading;</td>
<td>To search;</td>
</tr>
<tr>
<td>Scanning;</td>
<td>For research;</td>
</tr>
<tr>
<td>Serial/Non-Serial;</td>
<td>To summarize;</td>
</tr>
<tr>
<td>Reading;</td>
<td>For Discussion;</td>
</tr>
<tr>
<td>Single/Repeated Reading</td>
<td>Proof-reading;</td>
</tr>
<tr>
<td></td>
<td>While writing from multiple sources;</td>
</tr>
<tr>
<td></td>
<td>For text revision;</td>
</tr>
<tr>
<td></td>
<td>For critical review;</td>
</tr>
<tr>
<td></td>
<td>To apply;</td>
</tr>
<tr>
<td></td>
<td>For problem-solving;</td>
</tr>
<tr>
<td></td>
<td>For enjoyment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How a text is read</th>
<th>Why a text is read</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skimming;</td>
<td>Targeted reading;</td>
</tr>
<tr>
<td>Squirreling &amp; bouncing</td>
<td>Reading for class;</td>
</tr>
</tbody>
</table>

**Skimming**: this rapid reading style is used to quickly make sense of the overall argument of a document. Students tend to highlight phrases and sentences as a *mark of attention* or *make general comments* on the article, addressing the main points of the text. Meghan and Sophie pointed to the necessity of skimming articles due to the high volume of sources they had to read for their respective projects. After quickly reading the introduction and results section, both participants quickly summarized the article and made general statements about the content, deciding to keep or discard the source. While this is still in line with O’Hara’s understanding of the skimming process, we however augment his definition by adding that this mode of interaction appears to have greatly gained in popularity in this last decade due to the availability of sources online and in digital format.

**Squirreling & Bouncing**: while these online information-seeking behaviors are not necessarily reading types on their own, they are however linked to a skim-type of reading performed across resources. Chris and Mark talked at length about their seeking and reading habits online, mainly in terms of sources found on webpages. Both kept interesting web articles opened in separate tabs and differentiated between themes, topics and projects by using different browser windows. The browser technology however hindered their annotation strategies: both would *copy and paste interesting quotes* and passages to an external word processor file as a workaround to the browser. Other information artifacts resulting from these browsing sessions were generally
bookmarks, often through third-party clients such as Delicious.com, or bibliographic information, often copied and pasted to a word processor or text editor.

Other participants such as Meghan explained how they would download as many interesting articles as they could after a quick online skim, and then proceeded to read most of them in the same working session. Often these articles were related by theme and found through footnotes chaining.

These results confirm the findings from the different CIBER studies (Nicholas, Rowlands, Clark, Huntington, Jamali & Ollé, 2008; Nicholas, et al., 2008; Rowlands, et al., 2008). Our findings therefore demonstrate that the behaviors associated with the “Google Generation” (Rowlands et al., 2008) are already well in place in the daily activities of young researchers. While our participants do not match the description of the Google Generation (i.e. individuals born after 1993), they all are early users of the Internet and the web, most of them since their early teens. Since the CIBER studies called for a complete overhaul of the current digital libraries to meet the requirements of the researchers of the future (i.e., the Google Generation entering the workforce), we can conclude that there is currently a significant mismatch between the rudimentary digital infrastructures in place and the needs of our participants. The different workarounds that participants created (mainly situated around the browser and its inherent inability to allow annotation) confirm this assumption.

**Targeted Reading**: this type of reading occurs when one is seeking precise information from a text. Targeted reading involves scanning the article for keywords or desired themes. Students highlight or copy and paste to an external file the specific passages containing the relevant information. This mode of reading entails that a single text can be read multiple times throughout the research project, each time with a different purpose or target in mind. On this topic, Chris mentioned:

*I virtually never read the whole article... I will at most read half or two thirds of it, but I think that would depend on what I was using the articles for. For the projects I’ve been doing, I don’t usually read the whole article, and then I will intersperse it, like I will be reading parts during the research process, during the writing, during everything.*
**Reading for Class:** reading for class is a sub-category of targeted reading. Reading in a class context is often guided by the themes and concepts contained in the syllabus and the overall curriculum. Chris, a master student, noted how he always wrote down questions in the margins and reported comments and thoughts on a separated piece of paper when reading for class. This annotation strategy differed from the one he deployed when reading for his research project. These questions and comments were primarily designed to contribute to the in-class discussion, and thus acted both as a reminder of the content of the article, and as a contextual vehicle for further in-class discussion and questioning.

Reading for class is then inherently different than reading for research. Different institutional ecologies support and perpetuate these two activities. Reading for class involves different roles and contexts than reading for research. On the one hand, the annotator is a student, taught by an instructor in a class setting, using pre-determined goals and readings. On the other, the annotator is a researcher, guided by a supervisor, and dealing with open-ended goals guiding a more ad hoc review of the literature. Both of these ecologies intersect and feed off of each other, often creating tensions and negotiations between the individual and her annotating practice. These institutional ecologies and their inherent tensions are further explored in our discussion, section 3.4.2 *Annotation and Note-Taking as Materialities of Infrastructure.*

**Background Reading and Reading Around:** The bulk of the reading performed by graduate students involves reading around different topics, authors, theories, and ideas to gain an overall comprehension of their research topic. Students often pre-curate and check source material for potential suitability before engaging in this type of reading. Hence, activities such as skimming, bouncing around and squirreling are predecessors to reading around.

Reading around involves different types of annotation and note-taking strategies, ranging from highlighting as a means to sustain attention, to exhaustive commentaries on an article or book. Students annotate or take notes of material deemed important or interesting enough for future usage in the drafting process. The decision process leading to the selection or annotation of a passage is an interesting challenge for students. Indeed, students tend to keep their future needs in mind when annotating or keeping notes. However, these future needs are often vague and imprecise, since the research object slowly reveals itself as the student progresses. Hence,
annotating for future usage is a difficult task that consequentially prompts the student to re-assess her notes and the source material when writing.

Annotating and taking notes can also be beneficial interruptions of the reading process. While our participants complained that annotations tended to distract them from the main reading task, they also agreed that these disruptions were sometimes a welcomed break, allowing them to think and digest the text under study.

**Reading for Deep Comprehension:** this type of reading is explicitly reflexive and entails a linear mode of reading, from start to finish. In order to foster a deep comprehension and to allow the source material to be absorbed, students re-structure their source materials to reflect their mental representation of the topic. For instance, Alex used highlights and brackets in the margins and assigned themes and keywords to paragraphs or newly-highlighted sections of his source materials (see figure 22).

In order to gain a deeper understanding of the material, students restrain their annotation. They therefore limit themselves to techniques that require low cognitive overhead such as highlighting and underlining. This stands in stark contrast to “reading around” where note-taking and commenting are perceived as beneficial disruptions, creating the mental breaks necessary to digest the material piece by piece.

![Figure 22: Example of Restructure and Labeling of a Text](image)
Reading “in depth” as termed by Chris, was perceived as dangerous for the research project and as a guilty pleasure that he can allow himself only rarely due to time constraints and impending deadlines. Reading for deep comprehension tends to be time-consuming and only yields few notes or annotations. While the student may gain a greater comprehension of the text, these two limitations affect the research project negatively.

### 3.3.1.2.3 Conducting Research with Human Subjects

While none of our participants were conducting research using human subjects at the time of interviews, a few of them however discussed their experiences in handling annotations and notes pertaining to this type of data in the course of previous research projects. Students interact with interview data by highlighting or copy and pasting important passages to other text files. Participants such as Chris and used general purpose and specialized software respectively for coding interviews: both tagged their interviews with keywords or themes. Our study fails to address the numerous types of annotation and note-taking strategies in this research activity. However, we feel that such tagging technique is a type of annotation pointing to the adaptability of the practice.

### 3.3.1.2.4 Analyzing

We situate analysis between reading and writing, where both activities overlap physically and conceptually. Individuals have accumulated a sufficient knowledge base via their previous readings and slowly begin to form their own argument and structure their thoughts. The types of notes recurring in this situation appear to be more structured and formal: fleshed out outlines, structured thoughts, and longer and more complete comments tend are created at this stage. Students also create more personal reminders and memos as they attempt to integrate and shape multiple sources into a final output, all while under the pressure of deadlines and academic objectives.

### 3.3.1.2.5 Writing and Dissemination

Writing and dissemination is a natural extension of the previous phase. Students impose even more structure to their thoughts in order to produce a final end product. While all participants reported close working relationships with their supervisors or instructors throughout the research process, this last phase increases this type of communication as the revision process intensifies.
The intensification of the revision process and of the communication among committee members results in more structured notes appended to the draft. These notes are also composed in a more formal and explicit manner to be understood by individuals other than the annotator. As the student advances through the revision and editing processes, proofreading and editing notes, as well as personal reminders, become more frequent, often bounded by the same document. Alex described this diversity of annotation types and cited the importance of differentiating between the text in itself and the comments appended to the text through various visual methods:

*I do notes to myself in this kind of text, I put them in comments, so I tend to avoid to mix the writing that’s intended for the reader with notes to myself. Just to not forget. Or if I write in, I highlight them in any color, so I know that this might be something I have to look up. So highlighting in my own written stuff most of the time means that I say something that I have to verify a citation is missing, it’s a comment...so it can be different things I guess.*

### 3.3.1.2.6 Notes on the Activities of Research and their Corollary Annotation Activities

Very few studies have addressed the overall role of annotation within the research project, from ideation to dissemination. Previous studies have investigated the type of annotation created while reading on paper (Brockman et al., 2001; Palmer & Neumann, 2002) or reading on screen (O’Hara & Sellen, 1997; Liu, 2005; Hillesund, 2010). Some studies have ventured further and took a closer look at how different types of reading may be mapped to different annotation strategies: this was the primary goal of O’Hara and his typology of reading goals (1996) and of the research on active reading (Adler, 1940a; Kopak & Chiang, 2009). Research on the use and creation of annotation at later stages of the research process is however sparse. On the one hand the *creation, refinement, and formalization* of notes and annotation into an outline are explored by Case (1991) in light of the writing process (on paper). On the other hand, Hillesund (2010) and O’Hara, Taylor, Newman & Sellen (2002) explored the *use* of annotations and notes during the writing process. While our research has certain findings in common with previous studies (especially around reading and writing), we also differ from these studies by addressing annotation globally in the research project. Our research reveals the presence of annotation and note-taking activities throughout the research project. Annotation is an important support activity performed for a variety of purpose using different visual strategies.
3.3.1.3 Creating Links Between Documents: Annotations as Linkages

Annotation and notes also act as links across multiple documents. These often take the form of short notes (e.g., “see page XX, Bardier, M. 2002” or “Quote Latour?”). This type of annotation showcases the importance of distinguishing between an annotation in the context of the source material (on the material itself) and a note made in a document away from the source material. Participants often made notes about a certain text both on the text itself and in a separate document, contributing to a type of information fragmentation. They however closed this gap between the two representations by adding a note in one of the two documents linking back to the other.

Linking between two representations of the same text (e.g., the journal article and the corresponding notes) greatly helps the student streamline her use of annotations by alleviating the superfluous navigation to and from various documents. This back-and-forth movement is time-consuming and exasperating when one cannot recall the context of a note or annotation. Creating links between documents is therefore necessary for students. Mark, speaking of his note-taking practice, discussed this back-and-forth movement happening across documents, mainly between his notes and the source material:

So these sometimes have a lot of depth, but other times they’re just sort of like the guide post of the organization of the paper, so I can kinda look at it and be like “Ok, it’s like this” and then I can go back to the paper and this sort of tells me where things are [...]. I’ll often say things like “check out page blah blah blah again for this” I’ll go back in there and see... “see page 4 to 6, lots of controversy here”. So then I’ll go back and I’ll look at page 4 to 6, and then there’ll be like the highlights and the notes about it, and then sometimes I add more in here [the notes file].

Students also create links across multiple documents, creating a sort of hypertext by manually adding associative annotations to source materials. This type of annotation is closely related to more visual forms of thinking such as Maxine’s conceptual map detailing the relationships between different concepts or authors. This type of annotation is not simply an artifact from the reading process. Rather, it points to the interaction between reading and analysis as students recognize patterns across the literature. Annotations and notes that act as links between different
documents are therefore material evidences of this interaction and explicitly bridge different sources and activities.

### 3.3.2 How is the Information Encoded?

The findings in this section revolve around the content of annotations and notes, often heavily correlated with their formats and visual characteristics. We first discuss the use of language in notes and annotations. These notes run the gamut of structure, understandability and formalism. Research notes tend to move from one end of the spectrum to the other as the research progresses. Short, telegraphic and informal notes are more frequent during acquainting activities, whereas more structured and explicit notes tend to appear towards the end of the research project.

Conversely, students resort to symbolic markings for multiple reasons. Selection devices, such as highlights, are often used as simple attention devices, designed to help the student focus her attention when reading lengthy articles. Due to their low cognitive costs and the possibility of creating them without breaking away from the main reading task, highlights and underlines are extremely versatile markings for students. Highlighting allows students to remain immersed in a text: this justifies and legitimates the guilty pleasure of reading certain books or articles in depth for certain participants.

Students combine several annotation strategies, often juxtaposing textual and symbolic marks, as well as using color and typographic treatments to emphasize certain ideas. While notes and annotations (including their formats and visual characteristics) are often prompted by the content of the source material and the ideas of the student, the physical characteristics of the source materials and of the different tools used to create annotation also influence the different encoding strategies available to the student. For instance, the rigid layout of a PDF file and the awkwardness of the tools included in PDF readers may only allow the student to highlight or underline.

#### 3.3.2.1 Text and Language

Text-based notes and annotations are the most common encoding methods for our participants, but perhaps not the preferred ones since more efforts are required for creating these notes rather than creating a symbolic marking. However, certain text-based notes may be very close to symbol-
based markings: a telegraphic NO!, or even the simple use of punctuation marks often blurs the divide between text and symbol.

The type of language in the annotations surveyed in this study run the gamut of structure and understandability. This spectrum ranges from terse, abrupt language to lengthy descriptions and analysis of a certain passages of text presented in complete sentences with close to perfect structure. The structure and formalism of the language shifts throughout the research project, as the student advances towards composition activities. For instance, quick jottings about the scope of their research project are more frequent in the ideation phase, whereas more complete and structured sentences are the earmark of the analysis and writing stages. Complete structure and understandability are also present in notes and annotations intended for users other than the student herself, such as committee members or proofreaders.

We also distinguish between two types of written content. On one hand, student created directly authored materials, these marks “intentionally written in an effort to record information” (Bernstein, Van Kleek, Karger, Schraefel, 2008, p. 24). Questions or comments typed in a word processor, thoughts written in the margin of a book and even certain symbolic markings such as asterisks or stars fall under this heading. On the other hand, students also copied and pasted materials: quotes or passages pasted into a text file, waiting for future integration in the final written product. Finally, certain symbolic markings sit between these categories, both indicating an authorial intention, yet adding little new information to the overall material. Highlights and underlines illustrate this in-between category very well.

Our findings concerning the use of language in annotations echo and augment the previous studies conducted by Oviannikov, Arbib and McNeill (1999) and Marshall (1998). Our participants use different levels of structure and formalism in their annotations, effectively mirroring the explicit/tacit and formal/informal dimensions of annotations as revealed by Marshall. The setting of our study however allows us to assess the slow shift in language as the research project advances. Whereas Oviannikov, Arbib and McNeill (1999) and Marshall (1998) sampled notes and annotations created during the course of a single activity (i.e., reading a textbook), sampling across the research project helped us map different annotation events along the formal/informal and explicit/tacit dimensions. As the project progresses, the type of language
generally parallels the shift from reading to writing activities, from more informal and tacit annotations to more formalized and semantically explicit.

Two variables foster this shift: the final outcome and the choice of tools. First, students are constrained to use more formalized and explicit language as they near writing activities: the annotations and notes are ultimately assembled into a coherent written output. Previously cryptic notes are worked and re-worked in situ (in their drafts) in order to make complete sentences. Certain participants literally copy and paste their notes into their drafts and “string along words” as Maxine revealed. Second, different tools and software promote different levels of language, from PDF readers supporting more telegraphic markings to word processors encouraging more structured language. The effects of tools and software selections on the annotation practice are further explored in the next section, 3.3.3 What are the Preferred Tools and Mediums to Annotate and Take Notes?

Finally, text-based notes can also overlap with symbol-based markings. For instance, typographic elements (such as an exclamation or question mark) carry both an explicitly semantic and symbolic meaning. These markings blur the divide between textual and symbolic, and oscillate between the formal/informal and explicit/tacit dimensions. They appear to be hybrids of both types. On the one hand, typographic marks or even short interjections convey the tacit meaning of the symbol and are easy to create without radically distracting the student from the reading task. On the other hand, the typographic mark carries more semantic depth than other symbols, due to their use in everyday language.

### 3.3.2.2 Symbolic Markings

Students create multiple types of markings other than text-based annotations and notes. Highlights and underlines tend to be the most frequent marking strategies for student, especially when apprehending a new text, skimming new source materials or reading for deep comprehension. Emphasis marks such as bars, asterisk or arrows are also frequent among symbolic markings.

These symbolic markings are either used on their own or as part of a larger annotative act, alongside a textual annotation. Both strategies allude to different types of usage and contexts. Participants create the former when skimming, scanning and reading in depth.
of circumstances lead the student to actively restrain her annotation strategy to the sole use of highlights and underlines. First, skimming and scanning entails a fast reading mode when students are looking for specific patterns (i.e., keywords, general concepts) or assessing the suitability of the article for their needs. Due to the high volume of information that the student needs to process during these activities, highlighting is necessary to either bring attention to the patterns or simply to sustain attention.

Conversely, reading in depth requires students to be immersed in the narrative created by the author. To remain in this state of immersion, students must select annotation strategies that can be performed concurrently with the reading task. As pointed out by Mangen (2008) and Hillesund (2010), resisting distractions is often a very difficult task for students. We augment their findings by positing that the act of actively restraining themselves to using only highlights and underlines points to a series of compromises and trade-offs that students have to make with themselves. Students have to decide between reading in depth with or without annotating (which, in itself, is a choice that calls upon other motifs, i.e. reading for pleasure, or reading for research), and eventually settle on the type of annotation to create. Chris, for instance, discussed his hesitation to read interesting documents in depth during the course of his research. Reading in depth was associated with reading for pleasure, being immersed in a narrative and working through the author’s argument by reading intensively. Reading in depth, while a guilty pleasure for Chris, was perceived as an inefficient and time-consuming activity. The compromise that Chris reached is one where he would allow himself to read in depth only if he could use the material for his research, which entailed the need to annotate or take notes. Highlighting therefore legitimized his in-depth reading and allowed him to remain immersed in the text.

Symbolic markings can also be part of a larger annotative structure termed “standard association” by Marshall (1998). Highlights and underlines play a supporting role in this type of annotation, anchoring the main content (or the “body” as we’ve seen in section 2.1.2 Annotation as an Associative Device) and delimiting the scope of the annotation. This annotation format is generally reserved for reading around. Compared to reading in depth where overall comprehension is achieved by reading intensively and immersively, notes and annotations created when reading around are shortcuts to interpretation and signpost interesting ideas that may need to be revisited.
Symbolic markings such as asterisks or arrows demarcate interesting or important ideas. These emphasis marks may be adapted and designed to carry very personal meanings for students. Contrary to highlights and underlines that delimitate a scope, emphasis marks signal the importance of a passage, a section, or an overall text. Moreover, their ease of creation allows the student to *visually and dynamically curate information* and source materials by creating clusters of markings. The size of the cluster generally parallels the level of importance of the information (i.e., a cluster of three stars next to a paragraph may denote a more important idea than a cluster of two stars). Symbolic markings run the gamut of universal understandability, from terse and idiosyncratic markings that only the annotator can decode, to universal symbols which meaning is available to the common observer (such as highlights, circles or underlines).

### 3.3.2.3 Mixed Markings

Textual and symbolic markings can be juxtaposed on the same page, or even be part of the same note or annotation. Indeed, the concurrent use of textual and symbol-based markings may be part of a single annotative act (i.e., part of a standard association). The different elements composing standard associations are generally created at the same time, over the course of a single annotative event. The use of multiple types of content in one annotative act was especially clear in Alex’s annotations. Alex would first use highlights to block out a section of text in a journal article and then proceed to add an explanatory note or header right beside the highlighted text in order to explain this selection (as we have seen in figure 22). A similar juxtaposition of symbol and text-based markings occur when students gloss certain terms. Participants such as Alex and Sophie, whose first languages were not English, highlighted complex words and wrote their translation or explanation in very close proximity to the highlighted word.
The creation of different types of marks on the same page can also occur throughout the research project, over the course of multiple sittings, as the student returns to the material and to her notes. Figure 23 shows a page that has been revisited multiple times. In some instances, a symbolic mark serves to augment a textual mark, emphasizing an important element or narrowing down a scope: a student highlighting important keywords in her notes or circling a highlighted phrase on the source material. In other instances, multiple concurrent markings,
either of the same type or using different encoding strategies, are created to clarify certain thoughts or annotations, or to add other ideas.

3.3.2.4 Color and Typographic Treatment

Color and typographic treatments are analyzed separately from other types of encoding for their auxiliary task in supporting other types of markings. Color and type treatments are added to certain marks in order to enhance or modify their meanings. In some cases, color is the fundamental way of expressing the mark. For instance, a highlight would be impossible to imagine without color. It is the very nature of a highlight (and of the highlighting pen) to lay a certain amount of color on the page, either on paper or on screen. However, one can have multiple colors of highlights, each with their own personal significance. In this instance, color takes a meaning of its own. The use and meaning of colors for certain students is however not explicitly defined. This was evident when Alex indicated that:

*Or if I write in, I highlight them in any color, so I know that this might be something I have to look up... I have red in there, I think it’s because I just don’t want to overwhelm myself with all the yellow stuff... I know that this might be a minor point...*

Clearly, in this example, Alex uses color to sustain his interest and attention in an otherwise long and monotonous reading task.

Similarly, modifying the typographic treatment such as varying the weight, the font family (i.e., bold or italic), or increasing the point size, may denote different meanings. These techniques are commonly applied to digital material, whether the source material itself (if editable), the digital notes made in an external document, or the digital notes made on the source text. When used on the source material, this type of encoding acts as a text selection device. Conversely, when used in external notes document (i.e., on personal thoughts and ideas, or on a text that is already extracted from the source material), typographic elements serve to amplify, augment and emphasize certain portions of the text.
3.3.2.5 Encoding and Context: Consequences of Spatial Constraints on Annotation Format

The previous sub-sections hint at the importance of the material context of annotation and how it may act as a constraint or a resource for formatting and interpreting different markings. Since most of our participants annotate only on screen, the differences in annotating the source material and taking notes in a separated document are further exacerbated by the affordances of the medium and the digital tools available. This issue is discussed at length in the next section 3.3.3 What are the Preferred Tools and Mediums to Annotate and Take Notes?. However, we need to first emphasize the visual-spatial qualities at play in this situation affecting the note-taking strategies of students.

Different types of source material attract different annotation strategies. The tools available for handling the source materials also influence the choice of certain strategies. For instance, students tend to primarily use highlights on PDF documents. Emphasis marks are generally limited due to the availability of marking symbols in the reading or annotating software: students bring attention to specific parts of the PDF by using built-in shapes such as arrows (see figure 24). Participants rarely use notes: most interfaces require the student to interrupt her reading task in order to select the note tool. This interruption, a consequence of the software interface, shapes the annotation strategy of the student.

Furthermore the rigidity of the layout of some PDF documents, specifically the commonly narrow margins, deters students from writing lengthy notes and comments on the source material. Hence, while digital applications and software implement tools thought to foster an active reading strategy, the complete opposite was true for students. The design and implementation of these tools was counter-productive and constrained the student to a more passive reading strategy in order to minimize the distractions.

The built-in constraints of some tools prompt the creation of separate text files giving the student more space to reflect. However, these external files, often created using text editors or word processors, also carry their own material constraints. Full malleability of the document (i.e., adding freeform notes and comments in the margins) is very limited. Aside from the copied and pasted quotes, and the written thoughts and ideas of the participants, students also use color (in
highlights) and typographical treatments as means to annotate or bring attention to specific portions of the student’s thoughts.

3.3.3 What are the Preferred Tools and Mediums to Annotate and Take Notes?

Students combine an impressive number of tools to annotate and take notes. While a vast majority of these tools serve a dual purpose (i.e., supporting reading and annotating), our study also unearths an important amount of auxiliary tools and software supporting corollary activities related to annotation and note-taking. This section furthermore discusses the shift to digital annotation, observing a significant move toward a mostly digital practice. This finding contrasts with previous studies of the print/digital divide asserting that print-based annotations are still the prevalent mode of interaction with the text (Sheikh, 2004; Liu, 2005; Tenopir, King, Edwards & Wu, 2009). Our findings however do not strictly contradict this previous research. Rather, our study hints at the slow acceptance of digital-based annotations. The current digital practice is however far from the dreams of fully integrated digital annotations that several HCI researchers of the nineties attempted to implement (e.g., Schilit, Golovchinsky & Price (1998) or Marshall, Price, Golovchinsky & Schilit (1999)). While most students claim to perform an entirely digital reading and annotating practice, our interview data and document collection show that material, paper-based practices are still present, but in very specific contexts. Furthermore, the students’
digital practice is in fact a bricolage of various workarounds, and is far from being the streamlined practice that was envisioned.

This review of the digital and paper-based tools used to create and manage annotations and notes sits between the zoomed in and zoomed out states of our analysis. This hints at the pivotal role of such tools in the practice of our participants and indeed sets the tone for the rest of this results and analysis section. In this section, we first zoom in on the characteristics of the tools and software used by our participants. This prompts us to then zoom out and examine the consequences of these tools on the overall practice, while assessing the physical/digital divide and the workarounds that constitute the practice.

3.3.3.1 Digital Tools

Students use a core set of digital tools and software in order to create annotations and notes. PDF readers, PDF annotators, text editors and word processors are among the most commonly used tools by our participants. We distinguish two specific sets of tools serving different purposes, often used in parallel to each other: a first set facilitating direct annotations on the source material, and a second cluster of tools supporting note-taking. Participants also mentioned frequently using auxiliary tools and software that, while not directly linked to the creation of annotation, played an important role in the management, organization and use of the information accumulated via notes and annotations.

3.3.3.1.1 Annotating Digital Source Material (PDF)

Highlighting source material appears to be the most frequent mode of annotation for our participants. In fact, students would generally read their source materials with the highlighting tool selected, instead of the default cursor. This is especially true of reading source documents in PDF format, the most frequent format found in the field. Students open the files either in a PDF reader (e.g., Apple’s Preview, Adobe Acrobat) or a PDF annotator (e.g., Skim) after downloading the material or scanning the text into a PDF format. While both types of software are similar, we distinguish between them based on the primary functions of the software as defined by the participants, which tends to parallel their intended purposes as described by the issuing companies.
Students carry very different attitudes and opinions about the software selected to carry out this specific task. However, the pervasive perception of these tools and software as lacking certain components or a certain intuition is a common refrain among our participants. PDF readers with built-in annotation tools, while used by all students, are especially deemed to “get in the way of the real work”, as Sophie mentioned. Despite that students consider all PDF-related software as flawed, most still perceive them as a necessary evil for scholarly work. On the one hand, some of the participants were seemingly indifferent about their software choice. However, with further probing from the interviewer, students soon shared their preferences, likes and dislikes about different tools. Meghan illustrates this by stating that she didn’t have a default software package for PDF annotation:

*I just double click on [the source file] ... I don’t usually... I don’t specify whatever it opens with. I just find, sometimes Preview’s easier because all I want to do is comments. I don’t want anything... like I just want my little sticky on the side. Sometimes it opens up in Adobe, but like I just want it simple, cause I just want the sticky... so I find it’s easier.*

The more technophile participants in our sample shared their radically negative opinion of PDF readers and revealed that they would normally opt for PDF annotators such as Skim. While Mark excitedly revealed his love for Skim, he pointed to a fundamental difference between the software types: “*Adobe sucks... compared to Skim? Skim’s the shit. Skim’s awesome. It does all the different annotation types; the searching and I don’t have to pay for it. It’s open source, I like Skim a lot. Skim is good.*”

The availability of different annotation tools directly relates to the perceived lack of structure and formalism in the design of the software. This is illustrated in figure 25 comparing the toolbars from Skim and Adobe Acrobat. While both toolbars contain multiple annotation tools, Skim’s toolbar dedicates most of its space to annotation tools, offers more annotation formats and allows easier access to options (such as highlighter colors) that are otherwise buried in other software.
PDF readers such as Apple Preview or Adobe Acrobat allow only for few types of markings to be created, present notes in a very restricted graphical format and place them in defined spatial areas (see figure 26). These constraints structure the reading and annotating strategies of the participants and force a certain formalization of one’s thoughts in the note creation process. The formalization of one’s note-taking strategy occurs in the structuring of the annotation process. In order to create a note, the student is required to leave the main text area to select the appropriate note tool either in the toolbar or under the appropriate menu item. The student then redirects her attention to the text, decides where to affix the note by clicking on the page, and finally writes her note in the pop-up window.

Figure 26: Notes Created Using the Note Tool in Preview (left) and Adobe Acrobat (Right)
These steps are extremely delimited and salient for the student. These steps foster a certain externalization of knowledge hindering the management of tacit information. Certain participants countered this imposed formalism by creating workarounds in the PDF reader software. For instance, Sophie drew multiple horizontal bars in Acrobat and placed them at different angles to create an asterisk, which was easier to create than using the software’s built-in shapes (see figure 27). Other participants opted for PDF annotators where the larger selection of annotation tools was perceived as allowing more freeform annotations. The unstructured and informal characteristics of PDF annotators (often linked to the streamlined interface and uncluttered toolbars) facilitate the annotation of source materials. The bulk of annotations is generally created during the first thorough encounter with the text, when “reading around”. Therefore, creating an annotation or a note should not fully distract the individual from the main task of reading. Less structure and formalism results in a smoother transition between tools and tasks, and allows for thoughts to be jotted down quickly due to the low cognitive barrier present in the design of the software.

From Mental Images to Mental Models

The results on creativity show that apparently simple and de-contextualized mental images carry meaning that is not entirely captured by their visual and spatial properties (de Vega & Marschark, 1996; de Vega et al., 1996). The meaning inherent in images may be at least partly expressible in language and other abstract, symbolic notations. Sensory modalities besides the visual may also contribute to these models. Current theories in cognitive science support such multimodal representations of knowledge (Barsalou, Simmons, Barbey, & Wilson, 2005; Ramachandran & Hubbard, 2001).

What is found true for the kind of simple depictions employed in psychological research (e.g. the well-studied duck-rabbit figure) should be even more valid for the much richer images in science. Images in science are not simple perceptual entities. Not just the sophisticated images of high-level science, but the images occurring in school science too carry a great deal of conceptual, abstract and often mathematical content. Quite unlike images in art, and much less so than images in design and

Figure 27: Asterisks Created by Placing Single Lines at Different Angles

3.3.3.1.2 Taking Notes Digitally

However, students do not only annotate on the original source document as part of their practice. Taking notes in separate documents while reading and annotating PDFs often occurs concurrently. The reason for taking notes in an external document is also linked to the perception of structure and formalism that PDF readers and annotators impose. While PDF readers allow users to append a note to a line of text or a paragraph with more or less precision (see figure 26),
and perhaps afford a greater contextualization of the annotation, taking notes in an external document grants the individual more space to flesh out their thoughts. It is specifically this lack of (forced) precision and contextualization that lets students create a totally different type of commenting, one that may be more removed from the source material, and instead turned towards interpretation and “bigger picture” reflections.

Text editors, word processors and dedicated outliners are among the most frequent note-taking tools encountered in the field. These tools offer more flexibility to the student. Participants combined multiple types of notes related to multiple source documents in one text file, interweaving quotes, comments and ideas in one place. Figure 28 shows one of Alex’s daily note files, where he would collect thoughts, comments, and interesting quotes encountered throughout the day.

Figure 28: Example of an Intermediary File Collating Notes, Quotes, Thoughts and Other Materials

Furthermore, text editing and word processing software allows for more structure to be added to one’s thoughts as the project progresses. Mark revealed that his initial thoughts and ideas contained in his text file eventually grew into a structured and fleshed out outline, ready for final composition. More on this type of document and inherent practice can be found in section 3.3.4.3 Transfer.
This shift towards more structure and formalism is also visible in the student’s file management. Sophie and Meghan’s file systems revealed an interesting pattern of tool use. Both participants initially jotted down ideas in text editors (such as Text Edit or NotePad) and moved on to word processors once they acquired enough raw materials for more elaborate thoughts. Figure 29 shows Sophie’s file system for her notes and drafts. On the left, her notes folder boasts a mix of TextEdit and Pages files, while her outlines and drafts (on the right) combine Word and Pages documents.

![Figure 29: File Types Differ According to Research Activity (Left: Mix of Pages and Text Edit Files for Notes; Right: Mix of Pages and Word Documents for Drafts)](image)

Here again, questions of formalism and structure come into play. Lightweight software such as text editors only offer a limited selection of options and features, making it easy to load the application, open a document and write a couple of sentences. Software such as Microsoft Word or Apple Pages are more complete and substantial and allow the creation of more structured thoughts via their additional features. This is also why Mark opted for taking notes in a dedicated note-taking software: the structure imposed by the software is crucial for Mark as he sought stability in form and format across documents:

*Like it’s just useful like that, if you’re just looking for like this sort of simple way of structuring data, so you know, I can do sub-levels and I can shrink levels that I*
don’t really care... yeah, what it does, as much as anything is... it codifies it into a standard... my information is always laid out the same way, some of my notes, when I’m just writing loose notes, in a meeting... sometimes they’re just useless after... they’re good, but not when they’re like this.

The student’s use of tool and software progresses alongside the research project. The tools used at specific moments in the research project often mirror the complexity and reflexivity of the student’s thought process. Early, incomplete ideas are jotted down quickly, either on paper or in a text editor, whereas more complete, structured thoughts find their way to word processors. The choice of the word processor for note-taking and constructing more elaborate thoughts is in line with the movement towards writing activities. Taking notes and writing in the same software alleviates the task of switching between multiple open applications (i.e., switching between PDF reader, note-taking software, text editor and word processor), which distracts the student from the task at hand.

3.3.3.1.3 Dealing with Online Sources: The Case of the Browser

Reading, managing and annotating full-text, HTML sources is often a headache for students who would rather prefer to download PDFs and read them offline, at their leisure. However, most of them are faced with HTML-only online resources in the course of their research projects. Observing and interviewing our participants on this topic revealed multiple issues for online digital annotations.

First, annotation of HTML content was virtually non-existent for our participants. While certain browsers allow users to add plug-ins to ease this task, none of our participants availed themselves of these applications. One of the reasons for this lack of online annotation is the inherent separation of the annotations from the sources. This fragmentation holds great implications in light of the impressive breakdown of formats used by students. Our small sample stated that the majority of their sources were in PDF format (either full-text or image) and the rest was divided between e-books, physical books, journal article photocopies, online articles and websites. They preferred keeping their sources and notes together, in physical (or digital) proximity, which downloaded PDFs and text files could afford. Figure 30 illustrates this well: Alex kept source material and notes in the same folder. He further named both types of
documents using the same title, resulting in the close visual proximity of the source material and note file in his file management system.

![File List](image)

**Figure 30: Source Material and Notes are Often Stored in Close Proximity.**

Students also face the same inability to create digital annotations on e-books found through their university library catalogue. While user interfaces vary from vendor to vendor, students agree that none of them are user friendly enough to even allow proper navigational support. Here again, annotating the source material is out of the question due to the awkwardness (or the unavailability) of the tools, and the inconvenient storage of annotation, which requires registration and authentication with individual vendors. Figure 31 illustrates these constraints on e-book annotations and showcases toolbars from three different vendors. While both ebrary and MyiLibrary (top and bottom, respectively) allow annotations, they also require students to first sign in. Similarly, Scholars Portals Books’ interface (middle) is more streamlined than its counterparts, but the service does not allow annotation: students can only copy and paste passages, or export a limited amount of pages to PDF.
Figure 31: Comparison of E-Books Interfaces

However, web sources and e-books provoke interesting behaviors and actions. Indeed, web resources deemed important enough prompted complex negotiations and trade-offs for our participants. For instance, Meghan and Sophie tended to avoid web resources. However, if an article piqued their interest and warranted further reading, both would copy and paste the article from the website to a Word document, to then read and annotate the article. Sophie furthermore noted that she would create a PDF out of the word processing file in order to annotate the source material directly in her preferred PDF annotator. We thus observed a tendency for some participants to create workarounds by defaulting back to tools and software that are prevalent in their otherwise normal annotation practice. However, this time-consuming workaround occurred only if participants deemed the material important enough for their project.
The browser provides students with supplemental affordances that, when used concurrently with the previous workarounds, augment and refine their annotation practice. Students such as Mark and Chris availed themselves of the capability of browsers such as Firefox to save tabs and windows arrangements. They developed a sophisticated organization system where different themes or topics had their own dedicated browser windows. In these windows, multiple tabs were opened, most of them referring to an external note file (as seen in the previous paragraph). Keeping such tabs opened and ready to be selected allow students quick access to contextual information.

The presence of articles in designated tabs and windows also adds another level of contextual information about the sources and situates them in the greater project scheme. Chris addressed this point:

I certainly always have ten plus articles opened as I’m writing for the bigger papers, and I will sometimes... I’ll move them around, like if there’s 5 or 6 opened in a window, a browser window, I’ll move them around, reorganize the tabs to reflect the new structure of my paper and then when I get to a section... so I guess by reorganizing, I’ve already thought that you know, this paper is relevant for this section, this when I get to that section

Hence, tabs are perceived as material representations of articles: moving them around in the browser allow for an in situ organization of thoughts and topics. Their arrangement in browser windows is similar to the construction of an outline detailing the topics and their specific order in the construction of a coherent argument.

3.3.3.1.4 Auxiliary Digital Tools

This last example points to the panoply of digital tools used not only to create, but also to manage, organize, and use notes and annotations. Annotating and note-taking practices are thus deeply intertwined with multiple corollary activities. Multi-purposes software blurs the boundaries between different practices (such as annotating and organizing one’s files, or annotating and keeping track of one’s bibliography). This is clearly evident with bibliographic software such as Mendeley, which not only creates bibliographic listings of source material, but also provides organizational tools, as well as reading and annotating features.
Interestingly, the default file manager (the Mac Finder in the case of our participants) appears to be the most frequently used auxiliary tool in relation to annotation. The Finder plays many roles for the student other than for traditional file management: visual to-do list manager, retrieval system and organizational device for highly specific purposes. For instance, participants revealed that as deadlines were fast approaching, most of their Finder windows were dedicated to the research project, using separate windows for different tasks. Students therefore adapt and re-appropriate auxiliary tools to better serve their specific annotating and note-taking needs. However, these workarounds do not only concern digital tools. Rather, students also create interesting workarounds with physical, paper-based tools.

3.3.3.2 Print-based vs. Screen: The Rise (and Acceptance) of Digital Annotation?

Contrary to recent research findings, and to our great surprise, the annotation and note-taking practices of our participants occurred mostly in digital format. In fact, very few of our participants noted using printed journal article or other print-based sources. A minority seldom printed out source material, and would do so only if the sources were deemed important enough to warrant a longer and deeper reading. This last point indicates a type of conflict or contradiction between our participant’s practice and their opinion of the practice. While participants considered print-based reading and annotating more suited for deeper comprehension, many of them generally kept to digital sources and only dealt with print-based sources when all other options ran out.

Interestingly, this shift to an all-digital practice occurred quite recently for most of the participants surveyed. They were thus eager to tell the researcher about the differences between both mediums and to talk at length about the advantages of digital. When asked why they decided to forego print-based sources and activities, participants cited the high cost of printing multiple articles, the mobility issues when carrying all their printed sources, the searchability of digital documents and finally the ease of annotation on screen. Meghan furthermore pointed out a second economic argument: as a doctoral student, her funding allowed her to buy a laptop. She discussed how owning a laptop helped her completely shift from print-based to digital annotation. Hence, a recent increase in revenue often precedes any shift from print-based to digital practices.
Furthermore, keeping sources and activities strictly digitally-based also eases the cognitive load necessary when switching from one workspace to another, from the screen to the desk. Indeed, Sophie and Jacob observed how keeping exclusively digital sources and working only from the screen streamlined their practice. Consequently, keeping all activities exclusively digitally-based minimizes the constant back and forth movement between the computer’s desktop and the actual physical desktop and relieves the student from re-focusing her attention on a completely different space and format. The necessity and material requirements of the final output also guided this choice, as Chris described:

*Printing is expensive, it’s partially that... partially I was comfortable enough in annotating on screen that I no longer had to print them, and I wanted all of the sources to be digital or offline, and since I was using a lot of sources that couldn’t be downloaded in the process, I had to stick with digital, which was fine, I mean that’s where I was going in the end.*

Participants’ opinions of print-based sources were often radically negative. Some of them suggested that their use of printed articles was mostly imposed by a third party (e.g., an instructor). Mark mentioned how he would “*die a little inside*” every time an instructor would supply the class with paper copies of articles. “*Why would you do that? Why would you make it harder for me to ever use this again?*” Mark asked sarcastically. Conversely, other participants would make a point to obtain or print out copies of journal articles related to their classes. This behavior doesn’t necessarily contradict the digital practice as described earlier: printing out articles related to classes allows student to compartmentalize their different tasks and scholarly activities and distinguishes between class work and research work. Moreover, reading for class appears to be as an entirely different type of reading from what we have seen earlier in section 3.3.1.2 *Unearthing Diversity*, and entails a considerably different institutional ecology (which we explore in the discussion, section 3.4.2.3 *Social and Disciplinary Norms*).

Indeed, while reading for class is close to a type of targeted reading, the context of use is diametrically opposed to the research context. Students are often asked to bring paper copies to class in order to refer to various sections of the articles during class discussions. Some professors and instructors also impose a type of social pressure on students who bring their computers to class. Participants reported the drastic attitudes of some instructors towards the use of laptop
computers in class. Instructors considered computers as socially inappropriate for a class setting, often due to the belief that students may be tempted to do other things than focus on the task at hand.

This last point, the social pressure influencing the choice of medium and tools, is not solely restricted to the classroom. However, using this perspective as a starting point, we can see how paper and print-based note-taking activities are still present, even if most students admit to a solely digital annotation practice. Print-based annotation and note-taking practices appear in social contexts where others are physically present: the computer seems to be reserved for personal use or for collaboration across a greater distance. Alex referred to this double standard when assessing his own use of both digital and print-based mediums. While his practices seemed to be primarily digital, he revealed that the social implications of face to face meetings deterred him from using his computer, especially when meeting with his supervisor: “everything that I’ve written is on computer, so far. The only notes that are hand written are my talks with my supervisor... they are the only exceptions because I don’t want to type”. This is also echoed by Mark who, although a fervent proponent of keeping everything digital, slowly revealed the necessity of more physical practices in meetings, both in taking notes for oneself, via quick jottings on loose sheets of paper or in a notebook, or when sharing ideas with his research groups, by brainstorming and archiving notes on the lab’s whiteboard.

3.3.3.3 Physical Tools

Moving away from social considerations, we then turn to the choice of paper and other print-based tools. According to our participants, paper is still the medium of choice for quick and dirty notes left to oneself, reminders of sorts, and to-do lists. This echoes the findings of Bernstein, Van Kleek, Karger and Schraefel (2008) considering the role of information scraps for knowledge workers. However, we extend their findings by situating this type of notes and their use in the overall research project.

The notes created using paper and/or tangible mediums (i.e., whiteboard) are however very different than the notes and annotations made digitally. Indeed, screen-based notes and annotations generally lean towards an active engagement with source materials and reveal a dynamic construction of thoughts. Conversely, paper and print-based practices tend to be opportunistic and primarily contain meta information regarding the research project (to-do lists,
list of books to pick up), or raw ideas and thoughts that need to be offloaded quickly. Notes created outside the computer have a much shorter lifespan, which parallels the type of information contained in these notes (i.e., normally ephemeral prompts that quickly reach the end of their usefulness). This is especially true for notes taken on scrap paper, loose leafs, sticky notes, and whiteboards, four tools supporting an easy discard process.

These tools play a transition role in the work process. Students jot down notes on one of these supports in order to temporarily store their thoughts. However, if the initial thought is deemed important enough, student would then re-transcribe these notes to a more permanent support (normally digital). These notes thus act as a bridge between the here and now of sudden thoughts, and the future, more concrete actions, behaviors and documents; between the extreme ephemerality of the mind and the perceived permanence of digital note-taking systems.

Of the tangible supports observed in the field, sticky notes, notebooks, calendars, agendas, and loose-leaf paper were among the most popular choices, combined with tools such as pencils, pens and highlighters. These supports, mediums and tools represent a wide spectrum of formats and physical characteristics, as well as a corresponding range in the structure and formalism imposed by these formats. For instance, the large blank space of scrap paper attract more than just the odd, random thought: students tend to fill their pages in an organic manner, often juxtaposing different notes side by side. A calendar or an agenda however, impose a more rigid structure via its grid-based format and limited space. Nonetheless, some students adapt these tools to their own needs and often subvert their primary function by playing with the format or by imposing a different type of content. This subversion normally occurs because of the urgent need to offload one’s memory on the closest and most accessible support possible.

While our research does not specifically expand on the annotation and note-taking practices that dwell on the information scraps side, we however achieve similar results as Michael Bernstein’s team in their impressive study of knowledge workers and personal information scrap (2008). Indeed, we note how paper and print-based supports and tools are often used for their high visibility and quick positioning affordances. Participants tended to write important notes to self and personal reminders on paper and place them literally in the way of their daily activities. For instance, Sophie placed sticky notes on the keyboard of her laptop at night, fully knowing that she normally reached for her laptop first thing in the morning. This preemptive reminding
technique is however not only encountered in the tangible, print-based world: several digital workarounds capitalize on the same principles. For instance, Meghan highlighted the file name of her documents that urgently needed to be assessed throughout the day, also copying them to relevant folders (see figure 32).

Figure 32: Highlighted Files

3.3.3.4 Devices, Supports, Tools and Diversity

Despite multiple complaints over the tools present in their daily routines, students successfully use a complex constellation of devices, tools, supports and software, each with their own purposes and functions within the annotation practice. Multiple digital devices are generally involved in the daily research-related activities of students: laptops, cellphones, smartphones, portable media player and digital assistants (such as the iPod Touch). While most of these devices have unique purposes, students often adapt different devices for the same purpose and consequently fragment crucial information. This fragmentation normally occurs when the student either downloads source materials to different devices or annotates the same articles using
different platforms and devices. A similar type of information fragmentation and doubling happens across media, when students first annotate or take notes on paper and then later re-transcribe them in digital format. Information fragmentation is not a necessarily conscious decision. It rather points to a practice constrained by the necessities of the moment, itself defined by the physical and mental contexts of the student.

Information fragmentation divides sources and notes and leads students to feverishly scramble to remember where certain notes are stored. Conversely, information fragmentation can also be necessary since different contexts of interaction (i.e., a couple of words written on a blank page vs. the same words, highlighted in the original text) can provide different value for the user.

3.3.4 Annotation and Note-Taking Lifecycle

Contrary to the majority of research focusing on the creation of annotation, our study aims to explore the overall creation, storage, management and use of annotations and notes. Allowing this general goal to guide our lines of questioning with our participants proved fruitful. Following in the footsteps of previous studies of information scraps (Lin, Lutters, Kim, 2004; Bernstein, Van Kleek, Karger, Schraefel, 2008), our study uncovers a general annotation and note-taking lifecycle. Approaching annotation in terms of lifecycle allows us to assess what the participants are doing with the annotated information, when and for what reasons.

In addition to assessing annotation and note-taking beyond the moment of creation, the lifecycle allows us to observe how annotation and note-taking practices react to various constraints. Our analysis of participant interviews and collected documents reveals different variables affecting the lifecycle of notes and annotations. The overall research phase and status of the project, the available time, both in terms of current and projected timelines, and the participant’s discipline and area of interest are among the variables having serious consequences and a deep influence on the various phases of the annotation lifecycle.
This section effectively zooms out of the formal characteristics of note-taking and annotation in order to assess these practices in the research context. We next describe the different phases of our proposed annotation and note-taking lifecycle, as seen in figure 33 below.\(^4\)

**Figure 33: Annotation and Note-Taking Lifecycle, Adapted from Lin, Lutters & Kim (2004)**

### 3.3.4.1 Trigger

Notes and annotations are triggered and prompted by different activities and situations. While these triggers can occur at any time during the research project, they normally tend to cluster around reading and analyzing activities, with a final outburst towards the end of the project (correlated with commenting and proofreading). In fact, trigger events happen whenever an individual feels the need to take notes, annotate or highlight any type of document. Therefore, triggers normally occur in the interaction between the student and different types of materials, either her source material or her own documents (e.g., drafts, work in progress files, etc).

---

\(^4\) It should be noted, once again, that this cycle was designed with an academic research project in mind, hence most of the steps are geared towards the student’s research completion. We highlight this point in order to distinguish our research from the broader, more general topic of note-taking, micronotes, information scraps, and everyday information practices undertaken with a personal (i.e. not professional) aim.
Triggers vary widely throughout the research project. However, our research reveals that they are generally linked to four distinct contexts: the medium, the activity, the material, and the self. First, the medium in itself may trigger annotation or note-taking. Participants such as Mark pointed to the transience of information found on the web: Mark normally copied and pasted interesting passages of web documents into a word processor, just in case the original document would be inaccessible in the future (a fear that was common amongst the participants).

The activity itself can be a trigger for certain students. Sophie and Jacob both highlighted passages of their source material while reading, not necessarily because they thought a certain passage was interesting, but rather as a means to maintain their attention by engaging physically with the text. In this case, it is the activity of reading in itself that triggers the markings. While these may only be attentional markings, they gain a scholarly purpose when the student attends to the Refer stage of the lifecycle and attempts to sort out what was read from what was not.

The third context, triggers related to the material, is the most common and widely studied in annotation research. However, these triggers do not solely concern the material in front of the student. They rather emerge out of the tensions between the interests and goals of the student and the document in front of her. Here, students feel the need to annotate or take notes because something in the text is deemed interesting (yielding general annotation or marginalia), necessary for their project (sourcing a quote that will be copied and pasted directly into their draft), inaccurate (during proofreading), incomprehensible (leading to glossing certain terms) or would warrant further explanations (commenting on their own draft, writing personal reminders). These events occur when students read source material or when they read their own documents, and this, throughout the research project.

Finally, the self context stands for any trigger that was generated by the person’s own thoughts, either reflecting on her project, or trying to extrapolate from numerous sources and ideas. This is directly related to the interpretive notes as we have discussed earlier in section 3.1.1.1 Common Types of Annotation in Research.

3.3.4.2 Capture

The second phase, the capture of information, pertains to the creation of annotations and to the encoding of the selected information. This phase is characterized by its complexity, especially in
terms of the multiple decisions that the student must rapidly make. When faced with a trigger event, students must select what should be captured, as well as the appropriate method for recording the information, and this, according to multiple constraints: the student’s own interests, the perceived importance and complexity of the information, the time available to make the marks, the support medium of the source document, and the available tools, devices and software. These constraints affect the annotation in itself: the final choice of visual or textual encoding, the type of language or marks to be used, and the potential selection or scope of the thought to be captured. This thesis already addresses multiple variables at play in the creation of annotation (see sections 3.3.1 to 3.3.3). However, we need to emphasize how this phase in the annotation and note-taking lifecycle is replete with tensions, negotiations and trade-offs, mainly due to the amount of constraints one must deal with.

In particular, multiple trade-offs occur between the quality and structure of annotation, the time available to create the note or marking, and the tools or devices used. Maxine mentioned how certain ideas just came to her, and how she had to quickly jot them down in order to preserve this thought. In this situation, she normally reached for any type of paper, support and writing implement around her, and used mostly abbreviations in her text to capture the general idea. The resulting notes were lacking in structure and were generally incomprehensible. Maxine then transferred (and translated) the content of her note to a more permanent medium if the content was deemed important enough. The tension is then situated between the need to quickly make a note or annotation and the need for structure and richness of representation, later captured by transferring the original information to another format. Issues and conditions of transfer will be explored in the next section. Maxine’s example demonstrates the opportunistic nature of annotation and note-taking, especially when one is in a situation not necessarily amenable to note-taking.

At a smaller scale, a similar trade-off situation happens when students only highlight their source material instead of taking notes. This points to the detrimental influence of the back-and-forth movement between documents and areas of the workspace, even if taking notes in a word processor file might help structure and flesh out one’s ideas. Shifting to an external document is time-consuming, affects the cognitive load of the student, and overall deters the student from the primary reading task. The student must hold in memory the content of her note while shifting from the source material to the text file, hoping to remember the extent of her thought.
Participants were highly aware of this trade-off and thus started by highlighting directly on the source material, and only transferred and translated these markings to an external file later. This same trade-off also affects and delays the interpretation of the source material. Students tend to only highlight what is perceived as important while fully knowing that they will need to revisit the article in order to get a firmer grasp on the ideas and concepts of the text. Here again, capture speed reigns over the structure and understandability of the note created and consequentially leads to the transfer phase of the lifecycle.

The most extreme example of how time and capture speed affect other variables in a series of trade-offs is found towards the end of the research project, when deadlines are looming. Alex curbed his note-taking and annotating activities: “the more it comes towards the end, it gets very... I tend to take less notes and just to integrate the material quickly because then kinda what I write is my note”. Sophie also revealed that she often needed very specific information when writing her final draft, typically to fill some conceptual gaps in her text: her information needs were precise and time-sensitive. In order to fulfill this need, she engaged in purposive reading and scanned different articles for the concepts and ideas needed. She refrained from any type of highlighting or note-taking, and instead quickly paraphrased the idea into her draft, or copied a quotation to her text. Hence, the capture of note-taking and annotation tends to be dismissed from the scholarly routine of the student due to time limitations towards the end of the project. The annotation is still triggered by the specific context of writing, but the capture of the information is however performed in the mind of the student and immediately used in her text. A visual representation of this shortened cycle can be seen in figure 34.
Figure 34: Shortening of the Cycle due to Time Limitations

Trade-offs and negotiations with the self not only affect the method of encoding (and thus the structure and quality of the content), but also the scope of the content to be encoded. These types of trade-offs are the results of the tensions between the interests and goals of the student, the chosen theme or assigned topic of the research, and the projected use of the information. When questioned about his heavily highlighted files, Mark replied that it was difficult to select just one or two phrases to highlight. This is a common occurrence in the early days of a research project, when the goals and intentions of the students are still being shaped. Students tend to highlight widely, as the projected use may still be blurry, or in order to ensure the overall comprehension of the text.

3.3.4.3 Transfer

Our interviews and document collection reveal an intermediate phase through which information is transferred and may be refined and re-interpreted. Here, transfer refers to the process of moving information from one support to another, while either preserving the content, or re-structuring and re-working the information to better suit the affordances of the new medium. The transfer process also plays a crucial role in the refinement of information and in the overall thinking and writing process of the students. This section explains the transfer phase in the lifecycle of annotation, points to the various contexts prompting such transfer, and links this
phase to the overall progress of one’s research, understanding transfer as one of the main forces behind idea and category refinement (Case, 1991, p. 63).

The transfer phase in this lifecycle pertains to the shifting of notes and annotations from one medium to another. This phase may also encompass a translation element, acting hand-in-hand with the transfer activity, where information shifts from a freeform format to a highly structured one, affecting the content of the annotation. This is exemplified by Maxine who often transferred her notes from her notebook to a digital text file and simultaneously translated her cryptic notes from her notebook to more structured phrases and sentences.

As seen in the previous section, the transfer of annotation can be prompted by the current state of the note resulting from the speed of capture, as well as the student’s desire for a richer representation. Hence, transfer is often a consequence of the original capture. We identify four contexts of varying information needs prompting students to transfer and translate their original capture.

**Mobility and Accessibility:** Students transfer their handwritten notes to digital files in order to have quick and easy access to all their notes, at all times. This was specifically true for our sample, since all of our participants made the bulk of their annotations and notes on screen, either on the source material or in a text file. Outlier elements, such as ideas on scrap paper, risked being forgotten, which hampered their integration with other digital material.

These needs for mobility and accessibility also entail the shift from a transient, ephemeral storing solution, to a more permanent one. Certain students like Sophie transferred their digital annotations on the source material to digital notes in a word processor file. Here again, accessibility is the main reason behind this transfer. The repetition of the information increased Sophie’s chance to “run into it again” during composition activities.

**Structure and Understandability:** Students tend to transcribe and translate handwritten notes and jottings that were created rapidly and that lack in structure and understandability. Abbreviated words are extended and sentences completed. Shifting to a medium or to a platform that allows greater structure and richness of information is a primary outcome of this type of transfer. This transfer also affects not only the core structure of the note itself, but also the overall structure of note-taking files. Students maintain files where notes from a wide variety of
provenance are collated together. Transferring information from a single note or a single source to a file that contains a multitude of sources and notes entails a certain structure, if not chronological (in the order the notes files were assessed), then by theme.

**Re-Interpretation**: Extending abbreviated words and forming complete sentences, as well as adding one’s notes under a thematic header may lean towards the more interpretative role of transfer. Chris mentioned how he collated his thoughts and notes by opening all concerned text files, and then summarizing, or at least refining some of the overarching ideas in a single new text file. This re-interpretation is an important step in making and re-making working files, especially drafts, and in working towards writing activities: students generally work with a master document to which they add their thoughts based on other note files. This points to two interesting attitudes that emerged during our interviews. Participants admitted that they avoided at all costs the contamination of the master document with raw notes and citations: transferring material to the master document thus entailed the streamlining of language and ideas in order to fit with the master file’s content. Conversely, participants also tended to keep their note-taking files as intact as possible and created new versions of the file in order to add or modify specific elements.

**From Work-In-Progress to Richer Representation**: Transfer here occurs when students are ready to shift from a work-in-progress document to a more complete representation, and when they require different tools or software to do so. This shift also entails the process of versioning draft documents throughout the writing process.

These last two contexts of transfer (re-interpretation and shifting the content to a richer representation) hint at the crucial role of the transfer and translation of information in the overall research project. The transfer phase then acts as a means to “move things along”, to refine ideas and integrate material in working documents, all the way to the final written end-product.

**3.3.4.4 Maintain**

The maintain phase can be understood as an alternative or a complement to the transfer phase. Termed as the organization stage of the information scraps lifecycle by Bernstein, Van Kleek, Karger and Schraefel (2008), it encompasses the storage, archival and organization of notes
waiting to be used. This phase, as well as the refer stage, is the most affected by the time constraints, often completely avoided when time runs out or deadlines loom.

Our participants tended to maintain or store their annotations during researching and reading activities, therefore reserving their notes for future usage. Notes and annotation are stored until they reach a critical mass (i.e., until the student feels she has enough data and resources to start writing), or until time pressures the individual to start rounding up the project. This critical mass is however flexible. Alex described storing as much processed information or notes as was necessary to “get the framing right”. Once he felt he had the basic building blocks necessary, he then launched into a first rough draft. This rough draft, in turn, revealed the gaps in his research and directed his further information-seeking, reading and note-taking strategy. Alex went back and forth between reading and writing, often storing his annotation and notes only temporarily, before integrating them into his draft. Other students such as Mark felt they needed as much of the potential content as possible before drafting anything. Mark stored his annotations and notes for a very long time. Mark encountered multiple sources and read far and wide for a research project that had yet to have a solidified timeline. This led to an ad hoc organization scheme, storing his sources and notes in the most convenient way at the time of access or creation.

This last example leads us to discuss the multiple maintenance and storing schemes at work for our participants. We must first distinguish between the storage of the annotations and notes from the storage of the files that contain them. The former implies two possibilities: storing the information on the source document itself (what we have called ‘annotations’ so far) or storing notes in an external document (whether a text file or a scrap of paper). We have discussed the implications of the choice of storage support in section 3.3.4.2 Capture. The latter is what is properly affected by the maintenance and organization stage: the files representing the annotations and notes. The organization of these files varied from one participant to another, each favoring a different style. Alex first stored his text in a dated folder, itself stored in a thematic folder (e.g., “Browsing in the Library”), stored in a meta-folder pertaining to a phase of his research (e.g., “Reading”) and finally kept in a dedicated “Thesis” folder. While his organizing scheme was different from all other surveyed students, we can however discern general patterns in his strategy that can be seen across our sample.
First, digital notes documents are consciously stored in close proximity to the original source document if available digitally (as previously seen in figure 30). Proximity here refers to the visual closeness of different files in the student’s file system. This proximity stems from the need for students to have the best of both worlds: while they want the flexibility of a separate text file containing their own extensive thoughts, they also require the possibility to re-contextualize their thoughts rapidly, by opening both documents and referring back and forth between them.

Second, project-specific sources and note files are often stored in folders pertaining to the project. The overall file structure of the main folder however varies tremendously. Students organize their files chronologically, thematically, or by the type of activity involved (e.g., “Reading”). Students, however, also deal with sources and notes shared across projects, classes and meetings. Certain sources and note files, such as seminal articles that are the foundation of multiple research projects and discussions, often have several purposes. Consequently, this leads to a rather ad hoc organizational strategy. The most common strategy involved copying the source document and any related note files over to the project folder, in order to keep all sources and notes together. The main intention behind these organizational schemes is to increase the accessibility of the documents in order to ease writing activities by regrouping all sources and notes to be used in a single general location (normally the project folder), and making the notes visible (i.e., storing them at the same level as the source documents).

Finally, the organizational and maintenance schemes evolve alongside the project. This is especially true for students who use thematic folders to organize their projects. As new themes emerge and gain considerable importance during reading and analytical activities, new folders and hierarchies are created in order to reflect the evolution of the project. Sophie told the researcher that she “started with three folders… three themes that I thought were the key to my project, and ended up with 10 thematic folders where I organized my readings and notes”. Making and re-making one’s organizational scheme and file hierarchy is also an interpretative act in itself. It helps the student focus on specific themes and topics, distill the essence of articles and notes (in order to properly classify a source document or a note file in the right folder), and progress towards the completion of the project.

Organizing source documents, annotations and notes is not only restrained to the local file hierarchy. Alex, Sophie and Jacob also used bibliographic software and backup services, thus
increasing the complexity of their maintenance schemes. While at first they tried to maintain a similar hierarchy and classification scheme across all platforms, all failed to successfully mirror their local organizational scheme as sources and notes multiplied, time marched on, and deadlines approached. While all three students maintained that it was crucial for them to have similar organizational schemes across platforms, their usage however indicated that they mainly relied on their local file environment, often using bibliographic software or backup services as an insurance policy on which they would rely if all else failed. This is indicative of the wider trend of storing the same information in multiple locations (akin to copying source documents and notes from seminal articles from project to project). This replication across platforms, tools and projects may lead to information fragmentation, but may also be beneficial and necessary for students who want to keep crucial information visible at all times. Hence, Meghan duplicated her source materials and note files, and classified them in as many thematic folders as see fit in order to ensure that she would come across this information during her writing activities.

An important means of storing and maintaining documentation, annotations and notes for students is through the use of catch-all files (or intermediary files, as we have seen earlier), or through the more structured aggregation found in outlines and other pre-drafting documents. Such files, whether intermediary documents collating multiple sources or document outlines, were present in all of our participants’ strategies. Certain participants concurrently maintained both source-specific note files and catch-all notes files collating all information encountered thus far on a certain topic or a certain author. These types of files were consistently updated as new material was encountered and interpreted. Contrary to source-specific note files, these intermediary files are few and far between, but frequently accessed and modified, and often labeled distinctively, making them highly visible to student.

3.3.4.5 Refer

Here, the refer stage encompasses three separate steps: recall, retrieve and interpret. The recall and retrieve activities are directly affected by how the project’s materials (annotations or notes) are maintained. Once information is recalled and retrieved, the student may finally interpret, use and integrate the information in her project. Figure 35 shows the expanded refer stage of the annotation lifecycle.
The recall and retrieve tasks are often prompted by the current writing task of the student, which, in turn, is triggered by the critical accretion of information as described in the previous section. Students use different recall and retrieve techniques, according to the type of annotations available, as well as according to the type of information necessary at the time of use. The time of use relates to the specific moment when a student needs to integrate information distilled from her readings with documents representing the overall research of the student, from early drafts to polished writings.

Recall: The information needs of the student at the time of drafting or writing prompt the recall task. This information need might be vague (i.e., “I need to write my introduction”), vague with a direction (i.e., “I need to write a literature review section on information literacy”) or very precise (i.e., “I need that quote on information literacy in the 18th century from the book I read Tuesday”). Recall, or prompt, is highly context-sensitive. The nature and format of these prompts depend on the situation of the student in terms of project completion, and specifically relies on the student’s progress in detailing and fleshing out her ideas and interpretation. A tension generally exists between what needs to be accomplished by the student (the task at hand in light of the overall progress of the project) and the information encountered and encoded in annotation and notes so far in the project, often held in the student’s memory. Recall, then, is the necessary identification of an information gap in the student’s own ideas that may be filled by re-accessing the notes and annotations previously created.
Retrieve: Retrieve naturally extends the recall task. The student first identifies an information need and a potential information resource necessary to fill the gap (i.e., recall). She then needs to bring the necessary information back to memory, in order to use it as part of her project. This entails the re-access of information such as annotations, notes, and source documents in order to obtain a visual reference of the potential content. The need to retrieve and re-access the annotations and notes also stems from the fact that students, as per academic regulations, are required to avoid plagiarism: information is retrieved in order to be cited correctly or at least, to be used accurately.

Students select and use different retrieval strategies, depending on the precision of the prompt and on the information needed at the time of writing. Retrieval strategies are not only affected by the progress of the student in her project, but also, and perhaps more importantly, by the original format annotations and notes, and the student’s maintenance scheme.

As part of their research process, students need to collate and synthesize the information from multiple sources on specific topics or sub-themes. This may be necessary for the drafting of parts of the literature review, or for portions of the analysis. If the topics or themes are deemed important enough, students create folders dedicated to them, easing the retrieval of information. However, when trying to find more precise information (such as a quote by a certain author) or trying to recall certain concepts related to themes unmapped by the maintenance system, students tend to use an approach close to orienteering. They first orient themselves towards folders potentially containing the necessary information, and then take smaller steps within these folders and look at potential files and their neighbors one by one (O’Day & Jeffries, 1993).

Certain participants revealed interesting retrieval strategies: one solution which appeared to be a direct digital translation of on-paper retrieval, and a second digital solution which exclusively used the searchability affordance of digital files. The former consists of using the visual characteristics of files and folders in a digital system, coupled with the ability to copy, paste, and rapidly displace any portion of information around the desktop. Participants availed themselves of the material affordances of information representations that happened to be mainly digital. These same affordances were also found in printed representations as discussed by Bernstein, Van Kleek, Karger and Schraefel (2008). Content that was deemed very important and that was expected to be re-visited by the participants, was often visually highlighted on their workspace.
(i.e., the Mac Operating System allows one to highlight file names in the Finder, as we have seen in figure 32 showing Meghan’s highlighted files), and placed in the way of their daily doings by being copied over to frequently visited directories.

The latter retrieval strategy, using the affordances of the digital file format, was also correlated to a more ad hoc mode of annotations and notes maintenance. This is mainly exemplified by Mark who kept his source documents and external note files in very vague and non-descriptive folders, often simply tying the files to various project folders. Hence, no thematic or chronological arrangement was given to the file system. While Mark’s colleagues or some external observers may perceive this system as random and messy, Mark had precise reasons for justifying such organizational scheme. First, having a very small folder hierarchy allowed Mark to quickly save documents and files to the appropriate project folder, without having to navigate through a complex file system. Mark delayed the curation and interpretation of his encountered information and saved every interesting article from different areas of scholarship in the same folder. In fact, Mark generally delayed such curation until the moment when an information need prompted him to retrieve the relevant information. Mark simply availed himself of the global search function of his operating system and searched for specific topics and themes, which allowed him to collect documents from many different projects and classes:

I got to find my readings... God, my readings are everywhere... This is the one kicker is... I basically find my readings by doing a search. If I want to find something on like “mixed methods” [enters “mixed methods” in the Finder search bar]... and then I know that it’s in [name of the course], week 2 notes, week 9 notes, week 5 notes, week 7 notes and so I can go like this Week 9, right, and it comes up with this, both the source and the notes... So, it’s a combination.

Here again, the use of intermediary files or outlines bridging source material and interpretation is especially useful for students in their retrieval strategies. Since students deal with these files frequently by always updating and maintaining their content, retrieving the file in itself is not so much of a problem. Retrieval, then, is displaced from the search of the external representation of knowledge that is the file or folder on the individual’s desktop, to the search within the intermediary file, within the generally large collection of notes and quotes. The organization of these files varies quite a bit: chronologically, thematically, or by authors. Students invariably end
up scrolling through the file in order to identify potentially useful information, or use the search function of the text file allowing them to jump from one instance of a selected keyword to another.

**Interpret:** In this last activity of the refer phase, the retrieved annotations and notes are checked against the student’s information need in order to determine their usefulness. This interpretation task requires a second round of curation, as the student selects what is useful, relevant or necessary from her notes and annotations, themselves externalities from a previous round of curation in the capture stage (see section 3.3.4.2 Capture).

This phase also involves double-checking the retrieved information against the student’s memory. Due to the high volume of content in some projects, memories of concepts and theories tend to be misattributed or partly forgotten. If the student’s memory of the information indeed corresponds to the retrieved information, the student may continue her interpretation of her notes and annotation, and move toward the integration of this information with other ideas. However, if there is a discrepancy between what is retrieved and what is in memory, the student may opt to discard the retrieved information and continue her search for suitable content within her annotation and notes, or start a new annotation lifecycle altogether. In this case, this discrepancy acts as the main trigger, beginning an entirely new annotation cycle. Alex discussed how he launched into a targeted search for precise information and possibly for new quotes when he found that his previously created notes were too vague for his needs, and slightly different than what he recalled.

Once the retrieved annotations and notes are checked for accuracy, the real interpretation work begins. Students must use the content of their notes and annotations and integrate it with other material, either other annotations or notes, their working drafts, or other ideas. The integration of such material to other ideas (in other terms the actual use of annotations and notes) is also guided by different constraints. For instance, the academic policies on plagiarism structure not only the creation, but perhaps more evidently shape the use of annotations and notes. Meghan often used her redundant highlight and note-taking strategy. She generally transcribed word for word in a text file what she highlighted in the original source document, sometimes adding her personal thoughts. She then used the information in her notes file to literally play with her ideas and theories and referred back to the original source file to confirm and ensure that her notes and use
of quotes were accurate (this back and forth between both documents is facilitated by her use of constant referrers in her notes file, using the annotation as linkages type such as “see page 35”). This idiosyncratic behavior hints at a wider pattern of use that we have identified among our participants. While all of our participants used external text files for note-taking, they still opened and accessed the original source material when using the notes and annotations in the course of their writing activities, often simply to verify their assumptions. Hence, writing or drafting a document was often an activity comprised of an intense back and forth between multiple documents and types of documents. Students tend to have multiple documents opened when writing; documents that represent raw materials (closer to the source material) and other documents, such as outlines and drafts, that represent interpreted material (closer to the final research product). The ratio between these two types of material constantly shifts during the writing process, as the student need to rely less on external sources towards the end the composition, and more on refining her own ideas.

3.3.4.6 Complete

The complete phase is used here in a similar fashion as Lin, Lutters and Kim (2004). This analytic construct denotes the end of the period of use of annotations and notes. While not concretely assessed and discussed by our participants, this phase was still implicitly present in the creation, use and archival of annotations that we observed in the field. Complete then indicates that an annotation has been transferred to another document, that a highlighted quote in the original source document has been successfully integrated in the student’s drafted material, or again that a note has been double-checked for accuracy, failed to live up to its expectations and has been discarded.

We note that complete indicates the end of one period of use. Indeed, contrary to studies of other types of notes, including micronotes and information scraps, annotations and notes created in the research context can have an extensible usefulness and may go through multiple periods of use, often completing the Recall-Retrieve-Interpret cycle more than once (see figure 36).
Figure 36: Re-Use Cycle of Annotations and Notes (with Expanded 'Refer' Stage)

3.3.4.7 Discard

While notes and annotations may be re-used over time for different purposes, their material representations may at some point be discarded entirely once their usefulness for the task at hand or the overall research project has been exhausted. Whereas most notes and annotations are kept and archived on the student’s personal computer, students may also opt to discard their material.

Materials of trivial importance, such as notes and annotations created rapidly for almost immediate use (e.g., short reminders jotted down on a piece of paper) are normally discarded as soon as their usefulness expires. This is especially true of reminders or other meta notes dealing with the project itself. Conversely, notes and annotations pertaining to the research content tend to be actively archived and maintained on the student’s computer. However, some material might be discarded among the bulk of the notes: students discard documents, either source files or external notes files, which are thought to not be of any future use.

3.3.4.8 Archive

Whereas we found some examples of notes and annotations discarded after final use, we encountered a broader, more common pattern among our participants. Students were digital packrats and kept almost everything that pertained to their research, although vowing to “clean up” a bit, in the future. The main reason for archiving and maintaining their notes and annotations is the perceived future usefulness of the content. This packrat tendency is especially
true for students working on large-scale projects. The content created or used for the research might be re-used for a wide variety of end products such as grant applications, thesis proposals, comprehensive examinations and literature reviews. This perceived usefulness is also often structured around the student’s disciplinary standards, the core ideas and notions of her field, as well as its associated authors. Mark echoes this finding when he mentions:

> I just copy and paste it out of there into another document... it’s not like my research is changing that much, like... if I’m quoting Scardemalia and Bereiter, 1998, I’m gonna be quoting them forever because they’re like a foundational author, and I’m like... whatever, I’ll just pull it out of my other bibliography... I’ve got dozens of them... I just search! There it is! I’m like “Oh! Ok!” I’ll just go to the reference list and peel all the ones that I’m using again, because you have to! It’s the nature of it, you’re always quoting the same five people in everything you write, you have to... it’s like if you don’t put, for us it’s Vygotsky in there and if you don’t put Bereiter, people are like “You haven’t read the core literature!”... “Ok, Fine! There! I have it!”

### 3.3.5 How Does the Research Process Affect the Nature of Annotation and Note-Taking?

While we weren’t able to follow our participants extensively, from the beginning to the end of a research project, our interview data and collected documents nevertheless allow us to assert that notes and annotations are created and used throughout the research process. Moreover, since our participants were all at different stages in their research, we were able to observe the impacts of the research process on annotation and note-taking strategies. First, different types of notes are created at specific moments in the project timeline. Second, while annotations and notes may be created and used at any times in the research process, dominant patterns of creation and use are found at key moments of the research. The following section reviews the consequences of the research process on annotation and note-taking strategies and suggests that perhaps it is actually the different annotation and note-taking strategies that allow students to progress in their research project. Furthermore, the different annotation and note-taking strategies allow us to delimitate different types of activities involved in a research project. We distinguish between acquainting, analytical and composition activities.
The following discussion will refer to figure 37 and figure 38\textsuperscript{5}. The former visually demonstrates when different types of notes (meta or research) are created (top graph) or used (bottom graph) in the research process. Figure 38 digs into the research notes category to reveal two subtypes: here again, we differentiate between the creation and usage of these notes throughout the research project.

\textsuperscript{5} While our research did not handle quantitative data, we were able to assess the frequency and volume of different types of notes by discussing with our participants and analyzing the documents collected. The following graphics are meant as a visual explanation only, and should not be assessed for their precision.
Figure 37: Volume of Meta and Research Notes Created (Top) and Used (Bottom), Throughout the Research Project

1. Acquainting Activities
2. Analytical Activities
3. Composition Activities
Figure 38: Volume of Reading and Personal Notes (2 Types of Research Notes) Created (Top) and Used (Bottom), Throughout the Research Project
3.3.5.1 Acquainting Activities: searching, retrieving, reading

The early days of a research project showcase defined patterns of creation and use of annotation in. These few early moments are centered on the familiarization of the student with the grounds of the research: reading and understanding the research brief (even if self-imposed), searching for background information, retrieving documents and source material from different providers and finally reading around the topic. Two patterns emerge in this early phase: the creation and immediate use of temporary meta notes, and the creation of annotations and notes to be maintained for future analysis and interpretative work.

The former involves the creation of notes helping students orient themselves within the new research project: some keywords and names jotted down on a scrap of paper, the URL of a good web resource saved in a text file, or perhaps even a list of books to get from the library. These notes have an almost immediate use, and are often created to quickly offload one’s memory in order to attend to a greater task (i.e., writing down call numbers in order to be able to continue searching the catalogue). The creation and the use of such notes generally lead the student towards other types of work and other activities. For instance, taking note of meta information about books or listing tasks to accomplish may be necessary in order to move onward with the research. Such notes clarify the road that the student will have to travel and help them ease into the information gathering stage of the research. Our interviews revealed that this type of notes predominates the early stages of the research project, slowly yielding to more formal research notes as the student starts to read around the topic.

Indeed, as seen on figure 37, the creation of research notes slowly increases early in the research project timeline. Here, research notes encompasses two subtypes: reading notes, and interpretive notes. Figure 39 summarizes this typology and integrates the types as previously described in section 3.3.1.1 Common Types of Annotation. Due to the early stage of the project, and perhaps the lack of global perspective on the research topic, our participants discussed dealing mainly with annotations and notes that stemmed from their readings. These often take the form of highlights and underlines as the student tries to grasp of the scope of the project.
Implications for the Annotation Lifecycle

Two patterns emerge at this stage of the research project (see figure 40). The use of meta notes at this point in the project (and indeed throughout the entire project) is almost immediate. The participants seemed to bypass any transfer or maintenance activities (figure 40, left). The second pattern (figure 40, right) involves the creation of reading notes. Interestingly, these notes are not put to use immediately. Rather, they are saved for later, until the critical mass of information is reached and the student begins the interpretative work. Hence, the general pattern for reading notes at this point is the capture and maintenance of notes and annotations.

3.3.5.2 Analytical Activities: Reading, Organizing, Interpreting, Outlining

As the student acquaints herself with the topic and the scope of her project, analytical activities, from reading for deep comprehension to outlining the final project, also intensify. While meta notes are still present, and will remain until the end of the project, the volume of research notes clearly increases. This is illustrated by figure 38, which shows the peak of creation of reading and interpretive notes at the end of the acquainting activities and during the analytical phase. Conversely, the use of such notes is laterally displaced towards the latter stages of the project.
An interesting conclusion from our interviews might be that most reading notes, when used, tend to lead to the creation of more interpretative notes such as outlines or early drafts. In this sense, reading notes serve as a means for the student to move from the source material towards the creation and development of interpretations.

The creation and perpetual use of catch-all or intermediary files typically increase during these analytical activities. Creating and using these files entail the transfer of information from various source documents to a single file, collating quotes, thoughts and ideas. This type of file is often created or updated alongside the creation of highlights in the original source file. This dual strategy fosters a quick access back to the contextual information found in the original source file. While a large part of the content of these files seems to be mainly reading notes, it is the constant classification and re-organization of this content that qualifies the file as an analytical tool for the student. While some of our participants mainly displayed their notes in the chronological order of the original source material, others like Mark and Sophie created thematic files with themes and sub-themes dividing up the reading notes. This early classification was crucial for both participants as they worked towards creating their own arguments and as they prepared to create an overall outline of their final writeup.

While reading around different topics and themes occupied a great amount of time throughout the research project, our participants eventually reached a saturation point, having collected enough material to really launch into interpretive activities. While most admitted that they kept reading throughout the project, we observed a certain shift in the proportions of activities once the student collected a certain critical mass of notes and annotations. Students indeed move to an organizing and structuring stage, often re-accessing their notes, collating notes and thoughts, and moving towards the creation of a usable outline. Notes and annotations transfer furthermore frequently occur in this stage: this is evidence of the student’s working and re-working of her notes. Several participants mentioned that they created new documents as soon as they updated or modified their existing files. A certain language of *contamination* appeared in interviews. This language points to the perceived necessity of keeping one’s documents clean (i.e., devoid of raw material) as one moves towards interpretation and original content creation.
Implications for the Annotation Lifecycle

Analytical activities revolve around the transfer part of the annotation lifecycle. In fact, while notes and annotations are still triggered and captured during these activities, the transfer and translation of raw materials to more structured and organized representations is a crucial step towards interpretation. Hence two lifecycle patterns emerge throughout these activities. The first involves the creation of intermediary files or the concurrent use of highlighting and note-taking strategies (figure 41, left), while the second involves the refinement of the annotations and note-taking material via cycles of organization, structuring and classifying (figure 41, right).

![Diagram of lifecycle of reading notes and personal notes](image)

**Figure 41: Dominant Lifecycles During Analytical Activities**

3.3.5.3 Composition Activities: Drafting, Writing

Composition activities are a direct continuation of their analytical predecessors and share a similar modus operandi. Through transfer activities and the integration of interpretive notes, students refine categories and concepts, and structure their thoughts and materials along several main axes. Composition activities, primarily drafting, not only continue the category refinement task, but also eventually allow the student to expand on these categories and flesh out her ideas. Hence, while reading and interpretive notes are still created during these activities, the drafting and writing activities primarily involve the use of research notes.

These activities often lead to the simultaneous use of annotations and notes from multiple sources, which requires the student to shift her attention between multiple files and documents, in a complex back-and-forth strategy. This behavior involves multiple types of documents. While students attempt to fill in drafting and writing documents with their thoughts, they call upon external notes documents, annotated or highlighted material, outlines, and other previously...
created documents. The student must re-access multiple documents, especially notes and annotations, in order to bring back to memory what was read, what were the main arguments, and what her thoughts on the topic were at that time. This reveals an interesting contradiction that is ultimately unavoidable: while students create annotations and notes by keeping their future needs in mind, they re-assess and use their notes by trying to remember their previous state of mind. Thus, creating an annotation or a note is ultimately a type of negotiation between the present and future selves. Conversely, recalling information about a specific note or annotation is, in fact, negotiating between the past and present selves and often leads to the need to re-access the original source material.

However, other reasons also push the student to recall source material, alongside the annotations and notes made of the same document. While previous analytical tasks are more geared towards establishing the large strokes of the work, composition activities seem to be more detail-oriented and require the student to verify her information against the source. This does not necessarily indicate a lack of trust in one’s notes, but rather signifies the student’s compliance with academic and disciplinary standards. Indeed, the need to re-access original sources mainly for the proper integration of a quote is always present. Students would not only go back to check for accuracy of the quotes, but also to re-evaluate the words in context and judge if their use of the quote is indeed in line with the author’s thoughts and not taken out of context. By using a dual annotation strategy throughout the reading and analytical activities, the student eases this contextual task. Our participants created relational notes within their research notes (e.g., “see page 49”). These relational notations led to a certain page in the original source. In order to refine this retrieval search, students then assessed the page and look for highlighted terms or phrases: these indicated content that may have made it into the external notes file.

Working from multiple documents also increases the student’s chance to encounter the unfortunate consequences of information fragmentation. Information fragmentation often occurs when students take notes or annotate the same text in multiple format (i.e., on paper, on her iPad, on her computer). This fragmentation of platform and medium is often caused by the student’s need to create annotations and notes in the most efficient way, at a precise time. If a text is long enough, and the reading task distributed over multiple days, the chances of information fragmentation augment. Hence, sometimes a note or annotation not only prompts the student to re-access the original source file, but also requires the use of several other related files in order to
reconstitute the textual puzzle that information fragmentation causes. While only a few cases of information fragmentation were reported in our study, this problem may persist and potentially be aggravated due to the increase in the interest and availability of multi-purpose tablet computers such as the Apple’s iPad. The issue, then, is situated around the need to centralize and unify these representations, and to subsequently integrate and synchronize annotations from every platform and medium.

Finally, in addition to the final integration of notes and annotations in drafts and other documents, the creation and use of meta notes also intensifies specifically during the writing process and after having completed a significant portion of the draft. These notes are often reminders to integrate specific quotes and to refer to particular concepts and arguments. Additionally, multiple notes to self tend to be created on the drafts: students bring attention to parts of the text that require to be re-written or completed, to potential grammatical mistakes, and to serious gaps in their reasoning. These are often close to proofreading notes, although created and applied in a less standardized fashion than proper proofreading notations. A final category of notes generally created during composition activities, but that unfortunately fell outside of the scope of our research, are the traces of communication with other individuals, most often with committee members or colleagues. These notes fall in a category of their own: while being about the project, and thus similar to meta notes, they also entail the presence of the other which greatly influences the content and structure of the note itself.

**Implications for Annotation Lifecycle**
Composition activities revolve around the transfer task in the lifecycle (see figure 42). In fact, notes and annotations undergo their final and ultimate transfer when moving from their current location to the integration in a draft or final writeup. Hence, unlike analytical activities, which also revolve around transfer, composition activities ultimately complete the cycle and lead the student to either archive or discard the original note or annotation.

Moreover, this final transfer has much more complex undertones: this final phase is indeed another curatorial move on the part of the student. The student ultimately will not be able to integrate all the material encountered during the previous analytical activities. Hence, the student must go through a final recall-retrieve-interpret cycle and assess the relevance of her notes and annotations. Additionally, once the notes and annotations have been re-worked, interpreted and
placed in the composition documents, the student must then work with the material in situ. Participants accomplished highly complex analytical and interpretive activities within the context of the draft and refined their ideas, weaving them into a coherent argument. Hence, while one would not be able to recognize previous notes or annotations at this point, the student is still polishing them by iterating and editing her drafts and other composition documents.

![Diagram of annotation lifecycle]

**Figure 42: Annotation Lifecycle During the Composition Activities**

As previously stated in section 3.3.4.2 *Capture*, the annotation lifecycle can be dramatically shortened during composition activities. Indeed, these activities often take place towards the end of the project and consequently students face the pressure from colleagues and supervisory committees. Time definitely acts as a constraint in this case and forces the student to simply capture and integrate the source material directly into her composition documents, handling the interpretation work either in her head, or by working the material in situ.

### 3.3.6 What are the Attitudes of Students Concerning Their Own Practice?

Participants spent a great deal of time reflecting on their own annotation and note-taking strategy. The activity in itself was not invisible or blackboxed for most of our participants. They were very conscious of the bumps and kinks in their choice of medium, tool, software and overall technique, but often opted to delay proper assessment and modifications of such elements
of their practice until after having completed a large project, or until they would have more free time.

Some participants felt a certain social pressure to be organized and to have a streamlined annotation strategy. This was often the case when students were working in teams and would have to intellectually deal with colleagues who had a slightly different technique. Meghan expressed her concerns about her annotation strategy and linked them back to a collaborative project:

But, um, some of them, like I need to go through and... like I was here, actually listening to my friend, and how she was saying like, “I do little annotations for every article”, you know? So it’s easy to go through...and I need to be doing something like that, and I’m not looking forward to it because some of them... if I’m reading on a PDF, I just make comments or notes, so I could skim through it and see, like, what I was thinking when I was reading it that way. But I haven’t done anything, like, formal, and I don’t have any organization really.

While Meghan indeed had a formal organization of her own (highlighting, using the notes tool in PDF readers and transcribing her notes in an external text file), she was still striving towards a better organization system. The same tension could be noticed throughout our interviews: while all our participants were in search of a better system and found multiple faults to their current system, the reality is that most of them had a system that truly worked and that helped them accomplish large projects.

Some of our participants noticed a clear refinement of their annotation and note-taking strategies over the years, and even over the course of a few projects. Mark pointed to his practice, which changed tremendously between his Master’s and doctoral studies. The higher level of education, and higher level of expectations, played a critical role in shifting methods and tools. However, Mark also attributed this shift to his encounter with other academic genres such as grants proposal, which had very differently defined objectives, goals and publics. Hence, Mark stated that careful planning and note-taking...

... never mattered until now. You just had to be able to write, and they were like “Congratulations, here’s a cookie and an A”. But now, you’re competing with
other people, so it matters, right? You’re not competing for “everyone gets an A” when you’re writing SSHRCs, when you’re writing grants... you’re actually competing with everybody else that is getting an A, so it matters.

Overall, shifting strategy happens unconsciously and more frequently than what our participants mentioned. Participants changed annotation and note-taking strategies multiple times during a single research project. These small shifts and modifications to one’s strategy seem to be the consequences of the student responding to her immediate needs, and are in line with the findings supporting diversity in annotation strategies as seen in section 3.3.3.4 Devices, supports, tools and diversity. The material to be annotated and the context of annotation, here, drive change.

3.4 Discussion

The previous section reveals the messiness and complexity of annotation and note-taking as everyday epistemic practices. However, these practices are also shaped by larger structural patterns and contexts. This underlying structure differentiates annotations and note-taking in the context of the research project from their everyday or course-related counterparts: the epistemic nature of this micro-practice allows students to create and construct their overall research object.

The following discussion addresses the three objectives of this research as set out in our introduction, and attempts to fulfill this task by explicitly answering our research questions. Hence, this discussion first looks at annotation and note-taking as epistemic practices and addresses the duality of annotation as a relational dynamic as well as a partial material representation of the research object. The relational dynamic, here, is defined as the relationship that the researcher or student develops with her research object (the larger topic or argument of the research), in order to overcome the subject-object differentiation. This differentiation occurs when research work is novel and nonroutine, partly mediated by partial material representations of the research object. These partial material representations (such as books, articles and annotations) are the concrete and tangible stand-ins for the research object, pieces of the larger puzzle pointing the researcher towards other clues and pushing the research forward. Assessing the notions of “relational dynamic” and “partial material representations” requires to first zoom in on the materiality and content of annotation and to then zoom out to understand how these partial representations and their externalities affect the relational dynamic of annotation in the wider research context.
We then examine annotation as “materialities of infrastructure” (Star, 2002). This second move helps us understand how the relational dynamic is shaped and structured by layers of controls and access. We thus explicitly address infrastructures and institutions as part of an ecological understanding of annotation. This understanding is crucial in order to elaborate new solutions that not only support annotation and note-taking, but that also foster their integration in the larger realm of the research project, supporting other scholarship practices.

Finally, our analysis of infrastructures leads to the examination of annotation and note-taking practices in light of the recent rise of digital textuality. We thus conclude by discussing how annotation and note-taking practices have potentially shifted over the years by adapting themselves not only to new technological, material and digital environments, but also to the concurrent social and institutional changes that responded to (and helped create) these new environments. We therefore conclude by suggesting new approaches and by giving recommendations that seriously take into consideration the interplay between technological and social changes.

3.4.1 Annotation and Note-Taking as Epistemic Micro-Practices

Our results demonstrate that annotation and note-taking are at the core of the research activities of our participants and are indeed performed throughout the research project. In fact, they are stable across our sample, regardless of the discipline of our participants and of their theoretical inclinations. This hints at the “primitive” aspect of annotation and note-taking. Indeed, recent scholarship (Unsworth, 2000; Palmer & Cragin, 2008) alludes to the primitive function of such activity and understands it to be one of the basic building blocks of intellectual endeavors in an academic context. Annotating and note-taking are termed “scholarly primitives”, along with discovering, comparing, referring, sampling, illustrating and representing.

While we agree that annotation and note-taking are certainly a necessary component of “doing research”, we however take a more nuanced position on the concept of “scholarly primitive”. Our research shows that annotation can be described as a large set of activities, methods, constraints, and objects that may include smaller, more finer-grained activities. Therefore, this section looks at annotation and note-taking as micro-practices, a new term that allows us to explore primitive practices, themselves composed of scholarly primitives. We then move on to identify the epistemic nature of annotation and note-taking practices in the context of the
research project: the epistemic qualities of these practices distinguish them from their everyday counterparts. In doing so, we reveal that annotation is both the process by which the research object is created, as well as the externality of this process. Specifically, we highlight the bridging function of the annotation process, which defines the relational dynamic of the individual to the research, and creates externalities (the annotations and notes) that are partial representations of the research object. Characterizing the relational dynamic as a bridging mechanism and understanding annotations and notes as externalities of this process is crucial for research on annotation in two ways. First, it creates a renewed understanding of the materiality of annotation, and situates it as a necessary externality of larger processes such as reading and writing. Examining annotation and note-taking in an academic context cannot be separated from studying the activities that gave rise to the practice. Second, this sheds light on how the student may naturally move from one activity to another (e.g., from capturing an annotation to transferring it elsewhere) and even from one general practice to another (e.g., from reading to writing). Understanding this movement is crucial for creating future solutions and tools supporting not only the creation but also the use of annotations and facilitating their integration within wider practices.

3.4.1.1 Annotation and Note-Taking as Micro-Practices

As our literature review and results illustrate, annotation and note-taking are practices that students perform and enact over the course of a research project. This is the starting point to fulfill our first objective, which is to look at “what students do when they do research” (to paraphrase Miettinen, Samra-Fredericks and Yanow 2010), and more specifically what do they do when they create and use annotations in the course of their research project. Annotation and note-taking practices engage multiple dimensions as seen in the description of our practice framework. Our interviews and document collection therefore stressed on the methods, habits, activities, tools and objects necessary to enact this practice, as well as the relationship created between these elements as the context of annotation shifts throughout the research project. Our research therefore attempts to bridge the gap identified by Palmer and Cragin (2008) in information studies of scholarship practices: current understandings of the scholarly process lack a finer level of granularity, a level that requires a “holistic and materialist practice approach” in order to “understand the diversity of resources and activities involved in the scholarly process” (Palmer & Cragin, 2008, p. 170).
We therefore introduce the notion of “primitive” into our discussion, borrowing the term from Unsworth (2000) and Palmer and Cragin (2008). Scholarly primitives are simple tasks supporting larger activities, the building blocks of larger practices. We understand scholarly primitives as the tasks involving a direct interaction with the material of the research (e.g., the books, articles and drafts), and indeed emerging from the interaction between the student and the material. These aspects of scholarly work, such as comparing, referring, or sampling, often need to be combined with each other to form the basis of larger activities. For instance, our study reveals that writing from multiple sources involves referring not only to notes, but also to source documents, in order to compare and contrast different statements and arguments, sampling information from different sources in order to integrate them into a draft.

Most aspects of scholarly work, and specifically of annotations, have been assessed as isolated elements by previous research of broader scope. Human-computer interaction studies, and especially Cathy Marshall’s research, set the pace by looking at the materiality of annotation: the different marking methods and purposes of annotations, often in a textbook context. These studies are mainly centered on the tools that students use to create these marks on paper and aim to replicate the best practices in a digital context (Schilit, Golovchinsky & Price, 1998; Marshall, Price, Golovchinsky & Schilit, 1999). Other studies look at the function of annotations in different contexts: for students (Marshall, 1997; Ovsiannikov, Arbib & Mcneill, 1999), for knowledge workers (O’Hara & Sellen, 1997; O’Hara, Taylor, Newman & Sellen, 2002), and for historic readers (Blair 2004a; 2004b). Some studies attempt to link these functions to specific interpretive communities (Chartier, 1994; Blair, 2004a; Colclough, 2007) and to understand how the relationship between the book, the reader and the material interactions are structured by institutions (Nichols, 1991; Mayali, 1991; Slights, 2001).

Our approach differs greatly from this previous research on annotation, which atomize the practice. Our study uncovers a range of types and methods (sections 3.3.1 and 3.3.2), of tools and objects (section 3.3.3), and of activities involved in annotation (section 3.3.4). The components of the practice are tied to the research process (section 3.3.5), as we assess how their arrangement shifts throughout the research project. Perhaps, then, our research is more in line with studies of the research process and of reading and writing activities. Despite their orientation towards professional reading and writing, O’Hara, Taylor, Newman & Sellen (2002) addressed multiple dimensions of intellectual work and looked at the relationship between the
materiality of the activity, the spatial organization of the participants, and their repercussions on the wider tasks. O’Hara’s previous work regarding the different contexts of reading and the various reading goals of individuals (1996) can also be understood as a necessary predecessor to our research, pointing to different uses of support activities (such as annotation), of methods and of auxiliary activities involved in reading for different purposes.

Our results reveal how annotation and note-taking practices are intertwined with other research-related practices, and are often provoked by them. Specifically, reading, analyzing and writing seem to be three larger practices giving rise to annotation. Large surveys of academic professions, such as White (1975), Stone (1980, Case (1991) and Chu (1999) addressed the larger practice of what doing research entails for historians, literary critics, academic economists and humanities scholars in general. While their focus is mainly situated around finding key areas for improving library services for these researchers and involves a global look at the habits of researchers and the activities of research (with a lesser focus on objects and materialities), these studies help us situate annotation and note-taking practices as part of a larger ecosystem of practices.

Relating annotation practice to the concept of scholarly primitive demands that we understand the granularity and functions of primitives as posited by Unsworth (2000) and Palmer & Cragin (2008). Annotation and note-taking practices should be situated within larger practices, and simultaneously linked to finer-grained activities. Scholarly or information work primitives, then, are “‘self-understood’ functions [and] form the basis for higher-level scholarly projects, arguments, statements, interpretations” (Unsworth, 2000, online). We note that Unsworth’s intention in defining and listing different primitives is to uncover the “basis for a manageable but also useful tool-building enterprise” (Unsworth, 2000, online). We return to this point in section 3.4.3 Annotation and the Rise of Digital Textuality. Palmer & Cragin (2008), following in Unsworth’s footsteps, defined “information work primitives”, even finer-grained activities that are associated with larger scholarly processes such as reading and writing. Table 8 distinguishes between scholarly and information work primitives, according to both studies.
Table 8: Comparison of Scholarly and Information Work Primitives

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovering</td>
<td>Chaining</td>
</tr>
<tr>
<td><strong>Annotating</strong></td>
<td>Browsing</td>
</tr>
<tr>
<td>Comparing</td>
<td><strong>Annotation</strong></td>
</tr>
<tr>
<td>Referring</td>
<td>Collecting</td>
</tr>
<tr>
<td>Sampling</td>
<td>Searching</td>
</tr>
<tr>
<td>Illustrating</td>
<td></td>
</tr>
<tr>
<td>Representing</td>
<td></td>
</tr>
</tbody>
</table>

While our study demonstrates that annotation and note-taking indeed form the basis of higher level functions in the scholarly process, we concurrently discuss a multitude of finer-grained support activities giving rise to annotation and note-taking as a practice. In fact, our analysis reveal that most of the scholarly primitive functions listed by Unsworth are necessary for the global practice of annotation and note-taking. If we compare Unsworth’s primitives to our results, we distinguish two or more levels of granularity in his classification. Figure 43 illustrates our understanding of annotation and note-taking practices as being composed by some of the primitive activities revealed by Unsworth, and in turn situates annotation and note-taking practices within the wider realm of the research project. Therefore, annotation should not be strictly considered as a scholarly primitive. What Unsworth thought to be a single task appears to be a more complex set of activities, which necessitates the support of multiple single tasks. For instance, making a note in an external document involves *discovering* interesting information, *comparing* this information with knowledge contained either in someone’s head or on physical documents, opening a text file, *classifying* or *categorizing* the information and then finally *capturing* the note. However, we agree with Unsworth that annotation seems to be an activity common to all scholars and necessary to support wider practices.
Annotation, according to our results, is then what can be termed a *micro-practice*: a practice *supported* by scholarly primitives such as discovering, comparing and referring, but also *supporting* wider practices such as reading and writing. It is, in a sense, a *primitive practice*: a practice sustained by lower-leveled material-based tasks, while giving rise to and supporting larger processes. Hence, we expand on the meaning of primitive as discussed previously to encompass *sets* of tasks and activities that are irreducible in form (i.e., practices that are the
closest to the material of the research approached by performing finer grained activities and tasks).

These lower-level, finer grained activities tend to be closer to the material of the research, the material representations of the research object. These activities mediate the relationship between the representations (e.g., books, articles and notes) and larger processes (such as the creation of an annotation), as seen in figure 44. Furthermore, most of these tasks cannot be decomposed any further. These activities most often involve a single task (e.g., transferring a note from one place to another), and therefore must be combined with other primitive tasks in order to create or use an annotation. Conversely, annotation practice can be seen as supporting both reading and writing practices, indeed bridging the gap between these two larger practices. Annotation practice is therefore a primitive practice: a discrete set of finer-grained activities directly interacting with the material of the research and itself forming the basis of larger processes. It is one of the scholarly practices closest to the material and that boasts of a direct interaction with the partial material representations of the research. Conversely, larger processes such as writing may only access these material representations via different micro-practices, which mediate between lower and higher order thinking and materials.

The following section addresses the relational dynamic of the annotation process, the vector that fosters this movement towards the completion of the research project.
Annotation and note-taking practices in the context of a research project differ from their everyday or course-related counterparts due in part to their epistemic nature. Karin Knorr Cetina (2001) defined epistemic practice as “based upon a form of relationship that by nature of its dynamic transforms itself and the entities formed by the relationship” (p. 194). This stands in stark contrast to routine-based practices, where the object of the practice recedes in the background. Rather, epistemic practices occur when the researcher encounters nonroutine problems or new ideas. Knorr Cetina points to the dissociative dynamic present in non-procedural practices: the object of the epistemic practice is foregrounded and enhanced, which
prompts the researcher to overcome the subject-object divide by the deployment of relational resources. Epistemic practices are creative and constructive because of this constant relational dynamic that pushes the researcher forward, striving to overcome the divide. Hence, epistemic practices are characterized by the relational dynamic between the object and the researcher, the emotional valence rooted in this dynamic, their focus on epistemic objects, and the lateral branching out of the practice enabling the researcher to move towards other research goals.

**Relational Dynamic**

We first need to emphasize the importance of the relational dynamic for epistemic practices. This dynamic, here seen as the deployment of resources (emotional or functional) in order to overcome the subject-object differentiation, is at the core of annotation and note-taking practices as surveyed in our study and manifested in several situations. First, the act of doing research was relatively new for all of our participants. Therefore, every task and activity was far from being part of a well-established routine. This generates nonroutine problems and an increased reflexivity on the part of the student (i.e., How should I write my literature review? What is important to remember from this text?). The lack of knowledge concerning doing research then places the individual against the research as a whole: the act of doing research is foregrounded and very present for our participants.

This is also paralleled and augmented by our participants’ relatively new knowledge of their fields or disciplines. Thesis and dissertation projects help the student acquaint herself with the field. The student inevitably faces new concepts, authors and theories during the course of her research. We posit that this confrontation between the student’s current knowledge base and the new knowledge encountered in source material produces dissociative events where some (new) elements of the research object are foregrounded. These elements (e.g., concepts, theories, authors) stand out from the current knowledge base of the student and, as seen in section 3.3.4.1 Trigger, act as triggers for annotation. Hence, the relational dynamic of epistemic practices help explain the moments of creation of annotation, leading the student to move from a passive reading of materials to the inclusion of new information in her knowledge base, storing it for future usage.

Alternatively, the relational dynamic of this epistemic practice also fosters the use of annotation and therefore facilitates the move of the student towards writing and drafting activities. This
normally occurs when the student needs to collate all information gathered throughout the different research activities in order to present a self-contained output to her committee. The acts of comparing notes, referring to source materials and organizing one’s content bring the research object (the student’s main argument) to the fore. This was especially evident in the creation and use of outlines by our participants. Outlines were treated as the skeleton of the research paper, and showed the steps of the student’s reasoning. For instance, Maxine collated into her outline similar sources under large headings. These large categories were necessary to prove or disprove her argument. She then highlighted important quotes in this document that she absolutely needed to use, either to demonstrate her understanding of the field by citing a well-known author, or simply because a specific phrase summarized her thoughts perfectly. In either situation, Maxine consciously decided about what to use and what to leave out of her final draft. This curation process (as we have seen in the interpretation section of 3.3.4.5 Refer) required our students to step back and look at the overall research object in order to appropriately integrate different pieces of information into a coherent whole.

The Epistemic Object vs. its Partial Representations

The epistemic object, as defined by Knorr Cetina, builds on the notion of the epistemic thing as identified by Rheinberger. These objects of knowledge are diametrically opposed to their everyday counterparts in that they are at the centre of the research process: they are the objects of the investigation. Their open and question-generating nature points to a “lack of completeness” (Cetina, 2001, p. 190) that can never be entirely filled. In other words, the epistemic object is the overall topic or argument as defined by the student’s research question. While the student may achieve a pretty thorough understanding of her topic, and indeed answer her research questions, the research object itself may never be understood in its entirety due to its unfolding characteristic. For instance, a research project may be exploring the declining participation of women in computer science degrees. The research object, then, is the understanding of this decline and the exploration of reasons prompting women to enroll in different degrees. The finished report on this topic may give a very good overview of the situation, but will however not be the full representation of a completed research object. This stems from two reasons. First, the researcher herself may be aware of the faults and lacks of the research: the reasons or implications she failed to mention due to her lack of time, space or simply her lack of interest in exploring a certain area. Second, this report may even prompt more questions (as it should!) and
point to other areas of research necessary for a more complete understanding of the research object.

Since epistemic objects are constantly materially defined and can never be fully attained, the researcher can only deal with incomplete material stand-ins. These material representations can take many forms and always stand in relation to the whole (to the research object). For instance, to continue along our last example, the student researching the reasons for the decline of women in computer science may encounter multiple material representations of the research object. She might need to consult university websites in order to assess what is being offered and how the university is trying to attract women. She might move on to explore secondary sources discussing different topics helping her formulate her argument; she may opt to explore the types of discourses constructed in IT, the computer studies environment and the models of work and their filiations to a certain patriarchic schema. The books, articles and other material sources of documentation she encounters throughout this process are partial material representations of the research object, pointing to different facets of the problem. These representations help the student construct her research object by giving her small pieces of the wider puzzle. Concurrently, the representation points at what is missing: an article exploring gender in education from an institutional perspective may prompt the student to hunt for literature that explore the student’s perspective on this problem. These articles become stand-ins for the larger research object. Since the student does not have direct access to the research object, she may only deal with smaller partial representations of it: this is her only way of working towards a larger understanding of the object.

Hence, our participants dealt with and managed a series of partial representations: not only their source materials or their own drafts, but also their annotations and notes. All of these partial representations are parts of the whole: they suggest where and how to look further and give off clues that guide the researcher forward, towards the (unattainable) completeness of the research object. Concretely, our participants annotated or took note of interesting passages in their source materials because they recognized something that would advance their own work, that would fill a gap in their reasoning process, or that would simply prove their point: ideas, thoughts and comments that would get them closer to the research object.
Annotations and notes are partial representations of the research object in that they are pieces of the larger argument that the student is trying to construct. However, it is not only “what they are” that defines them. “What they are” and “what they are not” both define these representations, giving a glimpse of the research object, and pointing at what is needed to reach a more complete understanding of the research object. Their lack of completeness, as well as their relation to the overall research object, acts as a trigger for the student (as seen in section 3.3.4.1 Trigger, and in figure 45). Hence, the trigger element of the annotation lifecycle is a crucial phase where the lack inherent to the material representations assessed (the source material) prompts the student to create new partial representations, the annotations and notes.

![Figure 45: Annotation Lifecycle](image)

However, this trigger is not only due to the object itself. Rather, it points to a productive tension between the object, the representation, and the subject. Knorr Cetina points to the emotional valence of this tension partly created by the “structure of wanting” in the researcher, the desire of creating a complete representation of the research object by searching for new representations and moving on to them. This is the researcher’s attempt at bridging the gap between her and the object by deploying relational resources, as seen in the previous sub-section Relational Dynamic.

The tension between the structure of wanting and the lack of completeness of material representations, such as the source material (e.g., journal articles and books), triggers the
annotation and note-taking process. Hence, annotation and note-taking are not solely the externalities of the confrontation between the student and the material. They are also the relational dynamic that moves the research project along. It is through the annotation process that the individual tangibly relates to the overall research object. Hence, using theories of epistemic practices appears to be extremely fruitful when exploring the topic of annotation. Approaching annotation as both an epistemic process as well as the externalities of this process allows us to understand this practice as one of the main vectors pushing and pulling the student along, a dynamic rooted in the materiality of the research project.

**Transactional Engagements and Significant Symbols**

We can furthermore understand this relationship between the subject and the object as a transaction between the individual and the research object. We here build on and adapt the pragmatist definition of transaction as advanced by Mead and subsequently by Simpson (as seen in section 3.1.2.1 Pragmatism and the Practice Turn). The transaction at stake here is not between two individuals, as the classic definition supports. Rather, the transactional engagement is situated between the individual and the research object as a whole. The research object is used to describe not only the research topic or argument of the student, but also the disciplinary context from which it emerges. Indeed, the research object is a type of generalized other: the object embodies the general discourses, theories, concepts and expectations of the academic discipline.

This transactional engagement between the individual and the research object allows for meaning to be co-constructed by letting the researcher enter in a dialogue with the partial representation, mediated by significant symbols. Significant symbols are necessary for the student to assess partial representations: they act as the baseline for understanding, a knowledge base provoked by and reinforced by the partial research object. For instance, Sophie’s knowledge of her field of research was quite limited as she first started her research project. Her required foundational courses for her master’s degree allowed her to form a certain knowledge base: authors that have worked in a similar field as the one she was researching, key concepts of her discipline that she had to link to and finally a general idea of the current topics of interest for her field. These are the significant symbols that Sophie and her discipline shared in common. It is because of these significant symbols that Sophie may be able to effortlessly read and assess source material using similar significant symbols. However, as Sophie researched and read around her topic, she
inevitably encountered new concepts, notions and authors representing significant symbols not yet internalized by Sophie. This is when Sophie’s highlighting and note-taking practice intensified the most. In her words:

> whatever seems to be new to me and that I think will be good for my paper, I highlight. I won’t bother highlighting whatever I already know, I normally skip that anyways. I may make a note that the author talked about that in my notebook though. I also write names of people I don’t know in there.

Hence, significant symbols mediated the conversation between Sophie and the partial representations (books, articles, etc). Sophie assessed her source material by using her (limited) knowledge of significant symbols that she held in common with her discipline. However, if the source material contained information or symbols that were unusual for Sophie, she responded by highlighting or taking notes. We note that these new significant symbols, however new they are to our participants, are not completely out of the blue and unidentifiable. In her explanation, Sophie mentioned that she was able to gauge if this new information was important for her and for her project. This entails that Sophie’s own significant symbols played an important role in her attempt to decode new symbols and guided her interpretation and curation of information (as seen in Section 3.3.4.5 Refer).

This is not however a cold, emotionless relationship or transaction. Rather, this relationship subsists because of the emotional investment of the student in her research project. This is evident not only in the language used in some annotations created by our participants (i.e., an aggressive “NO!” in the margin for instance), but also in the frustrations that our participants described when they tried to manage and collate their notes and annotations for drafting purposes. The emotional valence present in this epistemic practice differentiates annotation and note-taking in the research process from their everyday counterparts.

**Annotation and Note-Taking as a Bridging Process**

The relational dynamic between our participants and the material representations then points to annotation and note-taking as a bridging process. Objects reinforce this structure of wanting by showing off what they lack. This prompts the student to move towards other representations and to slowly build her own representation of the research object by combining multiple partial representations. Thus, annotation and note-taking not only act as a bridge between previous
material representations (source materials) and the student’s understanding of the research object (via the Trigger and Capture phases), but also, and perhaps more explicitly, between reading and writing via the Transfer phase of the annotation lifecycle.

Participants transferred their notes and annotations to more structured files as their research progressed and as they amassed more and more content. Transferring is as an opportunity to streamline the content of annotations and notes, structure ideas and thoughts, and link different concepts together. In fact, the annotation cycle based on the transfer phase gives rise to the writing process, as students attempt to collect all the pieces of the puzzle into one coherent representation of the research object. This is exemplified by Maxine’s practice. She started her projects with a rough concept map of her research topic and slowly filled the gaps with notes and annotations as her research progressed. Maxine bridged acquainting activities with analytical activities by “mapping my key concepts, and then these are the researcher, and how... like I’m starting to make links between the concepts of the researchers and then where things can kinda congeal”. It is by visually drawing links between different areas of scholarship and fleshing out these links by adding notes and comments that Maxine moved towards analytical activities. When the concept map was not manageable anymore, Maxine structured the content and translated it to a text file. She continued to gather her notes in this document, under headings derived from her main concept map. Maxine collated all of her notes and thoughts into one document. This is a practice similar to Mark and Chris who copied and pasted passages and made notes into one single external text file. The result was a thorough representation of what Maxine gleaned throughout her research, supplemented by her own thoughts and structured in a way to ease her writing process. Indeed, as her research progressed, Maxine structured and restructured the contents of this file and created headings corresponding to crucial steps in her reasoning and in her argument. Therefore, the file facilitated the bridging process between analytical and writing activities. She explains: “I have my file open, and then... that would eventually... I would eventually just string words together. That would become that.” Hence, her notes and annotations stored in her notes file and structured over the course of the project acted as the raw material for her composition activities.

We note in passing that annotations, and the process that leads to their creation and use, bridge different temporalities in order to facilitate the movement towards the final use of annotations in the writing process. Indeed, annotations and notes are evidence of collaboration with the self,
attempting to determine what the future self will need in order to write a certain passage or trying to remember what the past self was trying to say in a cryptic note. This brings to mind Mead’s notion of temporality, where past and future are in the actions of the present (as explained in our discussion of Mead, section 3.1.2.1 Pragmatism and the Practice Turn). According to Mead, the past in this situation is the multitudes of social attitudes and significant symbols in a social setting, while the future is often anticipated and enacted. Approaching annotation with this definition of temporality is fruitful. Our study reveals an interesting contradiction in the annotation practice of our participants, as seen in section 3.3.5.3 Composition Activities. While students create annotations and notes by keeping their future needs in mind, they re-assess and use their notes by trying to remember their previous state of mind. Thus, creating an annotation is ultimately a negotiation between the present and future selves. This negotiation is however mediated by the past, as Mead understood it. In this situation, the past is the overall knowledge and significant symbols created by a discipline. Hence, the student assesses new significant symbols by projecting and predicting her future needs.

Conversely, recalling information about a specific note or annotation is a negotiation between the past and present selves, and often leads to the re-access of source materials. Alex discussed both of these temporal relationships:

*I think when I read I just highlight stuff that I found during reading and think “Oh that sounds kinda interesting”... stuff that I think I will use later on during writing, it's just... I try in the text file to capture the notion of why this paper seems to be interesting. So it seems to be interesting to support point A that I want to make... so I go back to that when I write point A, and I use then the yellow things to go over and like “What is the section that I found interesting and why?”*, and then try to grab things that I can cite and quote.

Annotation both in its creation and its use, is a bridging process that leads the student from assessing dispersed partial representations to creating her own, more complete (but still lacking) representation of the research object.
3.4.2 Annotation and Note-Taking as Materialities of Infrastructures

While annotation and notes may be considered as externalities of the epistemic process shaped by the interplay between the research object and the researcher’s structure of wantings, our study also reveals multiple external constraints influencing not only the material representations themselves, but also their process of creation and use. These multiple constraints hint at points of friction (and points of decision) between the goals of the individual and the infrastructure supporting the practice. We thus understand notes and annotations as materialities of infrastructure: material representations created by tools, applications and software packages embedding systems of classification and standardization, and their inherent political and social choices. Hence, the processes of creating and using annotations are means of reproducing, shifting and modifying the infrastructure.

Our conception of infrastructure is greatly indebted to the work of Susan Leigh Star and Geoff Bowker on infrastructures and systems of classification (Bowker, 1994; Bowker & Star, 1999; Star, 2002; Star & Bowker, 2006). Our focus on the constraints of the practice echoes their use of the term “negotiated compromise” (Bowker & Star, 1999, p. 34) as points of intersection between the individual, the community, the institutions, and the infrastructure that supports them. Thus, our “infrastructural inversion” (Bowker, 1994) occurs primarily by foregrounding the practical constraints identified in our participant interviews. Specifically, our interviews point to common constraints faced by our participants: the digital tools and software, the medium and formats, the role of the individual as student and researcher, and finally the social and disciplinary norms. These points of tensions allow us to look at how infrastructure standardizes the practice and at how individuals not only perpetuates the standardization, but also how they resist and conversely adapt their environment to their practice. This section discusses the technical and discursive standards at play in the annotation and note-taking practices of our participants, since “working infrastructures standardize both people and machines” (Star & Bowker, 2006, pp. 234-235).

While none of the studies surveyed for our literature review explicitly address the role of infrastructures in annotation and note-taking practices, some authors however pointed to the consequences of the “formal qualities of the book (Chartier, 1994) or the screen (O’Hara and Sellen, 1997) on reading and annotating. Our current research indicates that the assumptions
literally built into the book, the page, the screen and the digital text files are not the only ones affecting the process of creating and using annotation. Rather, closer look at the tools and mediums supporting these processes bring to light not only the technical assumptions, but also the political, ethical and social choices that are folded within them (Star & Bowker, 2006).

3.4.2.1 Digital Tools and the Myth of Active Reading

Our participants created their notes and annotations mainly on screen. However, centralizing this practice and confining (most of it) to the screen did not necessarily entail a smoother and more streamlined practice: our participants used a multitude of tools for annotation, often using and adapting reading or writing tools for their own annotating purposes. Taking a closer look at this assemblage of tools and at the tools themselves, yields interesting results: the assumptions built into these tools reveal consequential lacks that students overcome by juxtaposing different solutions.

The tools used for creating annotations and notes vary far and wide: from PDF readers and PDF annotators to text editors and word processors. More interestingly, students use at least two of these software packages for annotating and note-taking. The need to use an assemblage of software and software tools demonstrate the restricted view of annotation built into these tools, as well as a lack of the understanding of the trigger-capture process. In short, these software packages, and specifically the ones designed exclusively for supporting annotation, fail to support the active reading strategy necessary for students.

The active reading strategies uncovered in our study defy the traditional definition of the term by Adler (1940a; 1940b). Instead of a lone conversation with the author, through the use of annotations on the source material itself, active reading for our participants implied reading multiple documents around the same theme during the course of the same session, the use of highlighting, the need to copy and paste quotes word for word into more structured documents, and the need to collate and combine information by transferring notes and annotations to a single file. Hence, active reading is not simply about creating notes, but rather about the creation and almost immediate use of these annotations for analytical purposes: it bridges acquainting and analytical activities. For instance, Chris described his complex reading process as an ongoing analysis of his source material and as an attempt to manage his growing knowledge base. Chris’ practice, not unlike Maxine’s, was focused around the creation of external note files containing
all of his reading notes and ideas. While he normally highlighted in his PDF files, Chris also copied and pasted quotes to the external document. However, this was not done mindlessly. Rather, when discussing his strategy, he mentioned:

*I was trying to slot them in to the earlier subdivisions of themes, so I took a lot of the quotes and data, and looked at what themes would come out of that. So... there used to be a lot more quotes, actually... some random ideas I was thinking of at the time, all these things that I've never ended up talking about with quotes attached to them... I use a lot of spaces, blank spaces, like carriage return to keep ideas separated... a whole bunch of things that I never ended up using, notes to myself in capitals. I use it most of the way through. I don't like contaminating the master document with the work-in-progress kinda stuff... so I won't actually do kinda the proper writing in there, but I will do the organization here, and then refer back to it as I go and change things*

Here, active reading is not solely oriented towards reading and comprehending. Rather, we extend the definition of active reading in order to include a vector of motivation and purpose. Indeed, while our participants did “combine reading with critical thinking and learning” (Golovchinsky, nd), most of them also attended to an ongoing process of structure and organization that ultimately resulted in guideposts for future writing activities. Active reading is then the integration of the current text under study with the student’s expanding body of accumulated knowledge.

Software packages such as PDF readers and annotators do provide the students with some options for annotation, via the availability of different tools. However, these tools are highly compartmentalized and refer to a traditional or mythical view of annotation, almost directly translated from the printed world. In this view, an individual assesses one document at a time, reading with a highlighter in hand, and a ballpoint pen not too far from them. Students shift back and forth between writing implements, using their fingers and hands as convenient indexical or reference aids (O’Hara & Sellen, 1996). Tools that adapt this model of annotation to the screen fail to understand the pervasiveness of annotation and note-taking throughout the research project and indeed reproduce a *reading-based model of annotation* that seems to even dominate the contemporary literature on annotation. Indeed, our study clearly shows how annotation and
note-taking are processes by which the student orients herself towards writing, towards forming a complete argument from a wide range of resources, including books, articles and her own thoughts. This is in part due to the nature of annotation as an epistemic practice. Our participants often related the partial material representations under study to the bigger picture, to the research object itself. Establishing this relationship, either mentally or by using notes and annotations, fosters the student’s move towards the completion of the research object. However, this relationship is not rock solid: because of the unknown contained in the research object and its representations, one can only estimate how a certain passage may fit with the rest of the research. Hence, participants tended to highlight what “sounds kinda interesting” (to quote Chris), material stored for later comparison, integration and possible use.

The tools supplied in these software packages furthermore embed and encode a certain level of structure and formalism that the student must then face. This structure and formalism are found mainly in the interface of the software, the toolbars and buttons designated for annotation tools, and the different interaction scenarios built into the software. The default view, as well as the general interface of the software, foster a linear reading strategy by laying out the document page by page. Furthermore, the software normally maximizes the reading window, leading to the screen’s real estate being completely occupied by the reading or annotating application. Both of these aspects of the application are what “gets in the way of the real work” as Sophie termed it. These software packages seem to only foster the single act of annotating. However, the gesture of annotating is only a discrete task in a long series of activities that composes the annotation process, including referring, comparing, or writing down thoughts in a separate document (as seen in our earlier discussion, section 3.4.1.1 Annotations and Note-Taking as Micro-Practices). Hence, most of our participants juggled multiple applications at once, shifting and moving windows around, or tabbing to the next application in order to create and use their annotations.

The availability and layout of the tools and toolbars also distract the student from their primary task. The tools available (normally such basic items as highlighters, sticky notes, colored bars) are generally found in the toolbar above the text, and are unselected by default. To create an annotation, the student must first reach for the toolbar, click on the appropriate icon, and then bring back her attention to the text in order to select the appropriate position for the marking. Switching tools thus entails shifting one’s attention from the text to the toolbar and back, leaving the student hoping to easily re-find the place where she left off.
This process is highly structured and forces the student to formalize her thoughts and needs by selecting the appropriate format for her marking. This especially conflicts with the behavior displayed in the early days of the research project, when the student confronts new material and navigates through multiple texts. Our participants were not ready to commit to the use of an annotation tool such as the sticky note when reading around a topic, since this entailed externalizing and formalizing their thoughts, which were often imprecise or fuzzy at this stage. However, this may also explain the general tendency to only make use of the highlighting tool in PDF readers or annotators, and consequentially the need to re-open and reassess one’s markings when outlining or drafting.

The structure and formalism built into these applications, as well as their inclination towards annotating-as-reading, prompt the student to supplement her annotation and note-taking processes with software packages that are closer to the writing end of the spectrum. While text editors and word processors also involve a very linear note-taking strategy (with students starting to write at the top of the page and often taking notes in a chronological manner), both types of application allow the students to collate thoughts, quotes and ideas from multiple sources, and facilitate the analytical and composition activities. This strategy also alleviates the student’s need to shift back and forth between applications and application window, in order to view multiple documents at once. Instead, the student creates representations of source documents that can be easily compiled within the same text file, giving her access to multiple source materials in the same location. These representations, often a medley of quotes, notes and ideas, are modular and mobile. This format is essential for working and re-working material: the representations can be easily copied and pasted again to other documents, allowing for the transfer process to happen in a fairly easy manner.

Hence, the format and content of annotations and notes may be partly indebted to the type of tools used for creating these annotations. Students tend to avoid the structure and formalism of PDF readers and annotators by strictly highlighting content, while creating an impressive collage of notes and quotes in separate text files. This partly points to the annotation-as-reading assumption built into these software packages, a vestige of their early days as read-only PDF software mimicking the printed page.
3.4.2.2 Mediums, Formats and File types

Creating an assemblage of digital annotation tools is necessary for students who wish to keep their practice entirely digital. Participants cited a number of reasons for avoiding print-based activities, hinting at points of tension in their practice that were constantly negotiated. They pointed to the need for searchability, mobility, and the smooth integration of content as reasons for their digital-only practice. The same reasoning applies when dealing with the digital sources themselves: PDF formats are preferred over e-text or full-text HTML articles for their apparent stability and integration with desktop-based tools.

Participants confined their annotation practice to the digital space mostly and expressed their dislike for printed articles due to the lack of searchability and the inconvenience of moving around with piles of paper. These two reasons point to larger assumptions in their argument. The former demonstrates the tight coupling of annotation with other research activities and the need to streamline and reduce the interstices between each of them. The latter points to the nomadic status of students (mostly master’s student) and the effects of this mobility on annotation practice.

Annotating and note-taking are deeply intertwined with other research activities and create a complex network of tasks and shared resources. The externalities of one task (such as the downloaded PDFs found during a searching task) become the resources for another (such as skimming), which in turn possibly shares this resource with annotating. Our participants pointed to the need to keep each of these tasks in a digital environment to ease their workload. Transferring any part of this chain to a paper-based environment was perceived as time-consuming. In fact, dealing with printed documents adds more tasks and activities. These activities are normally focused on the translation of the material from one space to another, including the translation of material from digital to printed space by printing them, and the translation back by either digitally copying notes, or by using handwritten notes to draft in a word processor (printed to digital). This last activity is furthermore cumbersome due to the repartition of the work across workspaces: students move back and forth between the physical desktop and the computer desktop and shift their attention (and information) from one space to the other.
Perhaps Chris said it best: “I wanted all of the sources to be digital or offline, and since I was using a lot of sources that couldn’t be downloaded in the process, I had to stick with digital, which was fine, I mean that’s where I was going in the end.” Two major forces seem to drive this digital practice: the network of scholarly primitives as previously discussed (searching, downloading, skimming, etc.) and the need to submit a final output digitally. By annotating and taking notes digitally, students minimize the interstices between the different tasks and activities and remove the need for translating back and forth the information from the digital to the printed space. Furthermore, this helps students move towards achieving and completing the final stage of the research project: submitting a completed report. Writing and drafting are helped by transferring notes and annotations to the master file and by working with the notes in situ. Creating digital annotations therefore streamlines the process of use and allows for an easy conversion from the original file where the annotation is kept, to the master document.

Interestingly, the necessity of keeping all research activities in the digital space seems to stem from the increased availability of digital material and the constant growth of digital libraries. Accessing catalogues and digital libraries is the first interaction between the individual and the research object. The type of material that students browse and collect is mainly digital. This also points to a certain disciplinary influence on the availability of the material, and ultimately on the format and medium used for reading and annotating. Previous research on humanities scholars (Brockman, Neuman, Palmer, & December, 2001) demonstrates that their annotation practice tends to be a hybrid of paper and screen-based activities. The main reason behind this hybrid practice can be found in the type of research object itself: humanities scholars, such as historians (Duff & Johnson, 2002) or literary critics (Chu, 1999) require physical access to their primary source materials, often printed books and other documents. Notes and annotations created when assessing printed source materials tend to also be on paper. We note, however, that multiple digitization programs are currently under way, slowly shifting the practices of the humanities scholar to a mostly digital-based one.

The disciplinary influence on the use of specific formats and mediums is thus evident in our study. Most of our participants belonged to faculties reclaiming themselves of the social science tradition. Their research objects varied far and wide, but were all deemed relatively new, and gained academic interest only recently. Our participants studied different facets of recent technologies as applied to multiple aspects of everyday life. Consequently, the bulk of the
literature necessary for our participants was available digitally, in part impacting on the different research activities dealing with source material.

Additionally, mobility was of primary importance for our participants. This was especially true of the master’s students interviewed: most of them were nomads in their work habits, working from home, the library, the cafeteria or the lab. None of these students had a permanent office where they could keep documents, photocopies or notebooks laying around: they reconstructed their workspace everywhere they went. Paper, including notebooks and printed articles, was perceived as cumbersome, especially since they needed to carry multiple printed items at a time, adding up to the weight of their laptop computers and other accessories. Hence, our participants made the conscious decision to keep everything digitally, therefore lessening the number of items to carry around. This involves the unification and management of sources, annotations and drafts, but conversely allows for the persistent arrangement of the workspace.

The tendency to keep sources and files in close proximity is in part allowed by the digital formats available for download to the student. Digital libraries, vendors and individual publishers made the move towards the PDF format. Adopting the PDF format for digital distribution helps solve cost and pagination issues for publishers. Indeed, a PDF file for digital distribution can often be created in a matter of seconds: journal or book layouts are created using digital software that allows for easy preparation of PDF files. This alleviates the costs of transferring, adapting or re-designing the content using a different software, for a different output. Furthermore, preparing PDF files from the master layout document also allows for pagination and layout consistency between printed and digital formats, therefore avoiding the need to develop a concurrent citation practice for digital files.

Students tend to download and use PDF files whenever they are available. In fact, they often go out of their way to obtain a PDF copy of a crucial article or book, sometimes even creating their own PDF files from the source material. Indeed, some of our participants scanned and created a PDF out of interesting sections of books or journals only available on paper. Others copied and pasted text found online into a text file and either annotated it directly or created a PDF file. This last instance points to the general dislike of e-text and full-text HTML publications. While participants enjoyed the availability of inter-textual elements (such as linking between articles), this advantage soon receded over the fact that annotation and note-taking tools were not well
developed, and that these notes were generally stored in a separate location from the rest of their annotations. Instead, some of our participants created their own intertext manually, combining notes from different sources in the same text files, or explicitly pointing to other files and other resources in their notes.

3.4.2.3 Social and Disciplinary Norms

The student’s reasoning behind selecting the appropriate medium, file format, software and overall format of annotations is finally considered in light of the role the student plays in her research project. Our participants belonged to multiple social worlds, where they performed different roles as researchers, students and apprentices. While these social worlds overlap, they are also defined by different tasks, institutions, actors and generalized others, all of which influence the annotation practice of the individual. This section looks at the social and disciplinary norms that support the research project and shape the annotation and note-taking practice of graduate students.

Students such as Alex talked about a certain “social pressure” influencing the choice of mediums and tools in his annotation practice. This social pressure generally occurred in situations where other actors were involved, actors who played a specific role, as well as stood in for a generalized other. For instance, Alex mentioned how despite restraining his annotation and note-taking practice to the digital space, he normally took notes on paper when meeting with his research advisor. This situation demonstrates the overlap and the friction between two social worlds as seen in figure 46.

These social worlds share many similar elements and actors, for instance the student and his advisor are two recurrent figures (as seen in figure 46). However, these actors take on different roles, call upon different objects and embody different generalized others. Zooming out of the practice, we see that the figure of the “advisor” takes on a secondary role in the day-to-day activities of the student and recedes in the background. Students gather knowledge of this actor through informal meetings, by reading up on her background and generally through previous relationships that may have taken place (for instance, in a class setting). Students do not constantly interact with this mentor. Rather, they internalize a set of expectations that the advisor may have regarding the student’s research. The advisor is thus a mentor guiding the student and embodying disciplinary knowledge.
Conversely, the advisor plays a stronger and more present role during official meetings where he or she assesses progress made by the student (the circle to the right, on figure 46). In this instance, the advisor takes an authority position and embodies the academic institution (one of the generalized others of the situation) that not only grades the student, but that also asserts the validity of the research. Zooming in on the practice, the authority dynamic established between the advisor and the student greatly influences the materiality of the annotation strategies of students. For instance, Alex opted to leave his computer aside and took notes by hand during his official meetings with his advisor. Alex’s gesture might then have been out of respect for the advisor and the institution, partly mediated through the perception of the computer as a nuisance for establishing this type of relationship.

Indeed, the computer takes on different roles and attracts different perceptions from the actors in different social worlds. This is especially true of the note-taking and annotating practices of students in a class context. Our participants discussed how several instructors viewed the use of computers in class negatively, primarily because of the belief that students were distracted by e-
mails, chats or other activities, and thus failed to keep track in class. This stigma seems to occur normally in the presence of others (e.g., in a meeting with one’s supervisor, or in a class setting), which points to the social pressure generated in these social worlds by the different arrangements and roles of actors and objects.

A similar dynamic can be described between disciplinary norms and the content of the annotation created. Zooming out, we understand disciplinary norms as a set of knowledge established in an academic field, created by a bounded set of authors who write on themes and topics relevant to the field. Our participants’ research projects were therefore an attempt at speaking the language of their fields, using and referring to shared significant symbols (shared between them and the discipline), authors and concepts. Here again, the university, as symbolized by the student’s faculty, validates the research of the student and recognizes the membership of the student (and her research) to the wider academic field.

Zooming in, we see how notes and annotations are material consequences of the student’s internalization of these sets of disciplinary expectations. In order to gain membership to her academic discipline (and effectively submit a valid thesis or dissertation to her home faculty), the student must assess her source material by comparing its content to her internalized set of significant symbols as described and reinforced by her discipline. Finding a match between the source material and the internalized symbols may result in no annotation or notes whatsoever: the source material only contains common disciplinary knowledge. Rather, the student is on the lookout for material slightly diverging from this set of expectations: this material in turn becomes the trigger necessary for the annotation lifecycle to begin, as previously described in our discussion in section 3.4.1.2 Annotation and Note-Taking as Epistemic Practices.

3.4.3 Annotation and the Rise of Digital Textuality

The topic of annotation and note-taking benefitted from a sudden increase in academic interest in the mid to late nineties. Along with searching, browsing and reading, annotation and note-taking were among the scholarly activities that were expected to undergo drastic modifications due to the then-recent technological advances in libraries and publishing, and the development of new digital tools. In light of this imminent technological shift, scholars were prompt to observe and analyze how individuals went about doing their daily tasks, and how these tasks and activities could be translated and enhanced in a digital environment. Hence, a number of studies assessed

Annotation was then understood as a single task to be translated from paper to screen, and that could benefit from the affordances of the digital medium (such as searchability and shareability) along the way. Most of the digital annotation systems suggested in these studies dwelled on the potential hybridity of the practice: retaining the physical affordances of paper while augmenting the practice with new technologies. This conception of annotation was often translated in a literal way: pen-based computing garnered sustained attention from the HCI community up to this day (Marshall, Price, Golovchinsky, Schilit, 1999; Schilit, Golovchinsky, Price, 1998; Liao, Guimbretière, Hinckley, 2005). Pen and tablet-based computers are often regarded as the ultimate scholarly tool, mixing the flexibility, tangibility and ease of use of the traditional pen with a screen-based reading interface, providing searchability, hyperlink capacity (for finding related material) and multifaceted navigational support (from full screen to marked-up thumbnails).

These systems however failed to gain widespread adoption up to this day. While pen-based computing left the academic test bench and made its debut in the commercial realm (i.e., the Livescribe system using the SmartPen technology), we have yet to see a sharp increase in the use of this technology for students and scholars. This lack of success hints at multiple issues embedded in these systems, including the paper-based metaphor and the lack of integration with other scholarly activities.

In order to address these issues, we need to first assess annotation and note-taking as part of the larger context of digital textuality. Our definition of the term “digital textuality” is manifold, but essentially indebted to the digital humanities, a field concerned with the integration and use of new technologies in scholarship practices. Textuality refers to how a text is read: the medium, tools, interfaces and other characteristics that mediate the reading, reconciling the text and the reader. These features envelop and convey a text but, as we have seen throughout our study, are not self-generated. Chartier (1994) brought to light the social nature of textuality: publishers, printers and authors contribute to the physicality of the printed book. Likewise, we can
understand digital textuality as constructed by designers of e-book or online journal interfaces, publishers, vendors and even authors.

Hence, digital textuality is supported, maintained and developed by a delicate network of individuals, communities and institutions who produce and use not only digital texts, but also a plethora of publishing, distribution, searching and reading tools. Furthermore, digital textuality itself can be defined as a multifaceted entity. Just like the physicality of books differs from the one of magazines or pamphlets, and therefore entails different ways of reading, digital textuality has multiple disguises: digital text can be apprehended from a multitude of perspectives, via many different interfaces.

Our study confirms the frequent use of journal articles in PDF format. These files may be recent articles, directly exported into a PDF format from the original journal layout master file. This normally entails clear and highly legible typography, offers the possibility for publishers and authors to directly link in-text citations to the appropriate bibliographic notice at the end of the article, supports hyperlinks to documents on the web and allows the user to highlight or copy and paste text. These PDF files somewhat differ from those generated from older articles first published in print and recently digitized in order for vendors and libraries to offer easy access to the journal’s back catalogue.

Hence, producing a PDF file from an older article requires a different strategy and necessarily yields different results. For instance, the typography might be of lesser quality due to the scanning process, and sometimes might be un-selectable unless the student (or the vendor) runs an optical character recognition (OCR) software on each of the pages. Hence, these older articles do not necessarily support highlighting or other types of text selection methods: students often encounter several problems when copying and pasting text to external files due to the OCR software’s approximation of letters and typographic symbols. Therefore, even if the final output may be the same on the surface (a PDF file of a journal article), they both involve the use of different tools in their creation and use, and therefore imply a different digital textuality.

The rise of digital textuality as we state in our section title refers to the current state of the digital textuality in the research context: an entity that now has certain stable roots, such as sets of common interaction principles and interfaces, yet is still expanding and growing towards areas vaguely defined. Therefore, while PDF files of journal articles are well integrated into the
student’s mores, other types of digital content and digital textuality have yet to find widespread adoption. PDF files of journal articles play on a well-known paper-based metaphor: that of the printed journal publication. The format and layout of the pages are the same due to reasons seen in our previous discussion section, which may have eased the transition from print to screen for many students and scholars. The textuality, then, is similar to the physicality of the printed journal article, demanding a similar type of manipulation which is however constrained by the characteristics of the screen itself (i.e., I can easily jump from page 1 to page 34 in both printed and digital versions, but the very act of reading text on a screen entails scrolling down to view more, a gesture absent from the material world of the printed representation).

Conversely, our study demonstrates that students and the wider academic community have yet to completely adopt and fully learn how to deal with web resources. More specifically, we observed how our participants’ annotation and note-taking practices were stifled when facing material completely native to the web, such as web publications, online HTML journals, and all manners of blogs or wikis. This situation might also be soon exacerbated: as web-only resources increase in availability, and as traditional publications migrate to the web, students will be faced with a majority of resources that may only offer HTML text, foregoing PDF distribution altogether. Students will then be forced to adapt their annotation and note-taking practices to make the best out of these web resources.

Current activities dealing with web resources normally involve the copying and pasting of text to external, local documents (text files mainly), concurrently with keeping and organizing the original material in browser windows and tabs. However, this tendency to copy, paste and create local representations of web content may not be solely linked to the need of unifying all source material in one place, or even to the need of using familiar software to annotate. We believe that keeping local copies of material and notes is necessary in the writing process and that students will tend to download, copy, paste and construct local representations of online material as long as the writing process involves desktop-based software such as text editors or word processors.6

6 This, in turn, is also influenced by the demands of the faculties, schools, universities and publications asking for specific formats and layouts of documents, often requiring a word processor to create the output.
However, as web-only resources grow in importance and availability, it will be crucial for students to become familiar not only with new tools, but also with new sets of activities and tasks related to reading and annotating online. Several companies have already attempted to bridge this gap, providing tools to annotate content online in a persistent manner (i.e., keeping the annotations permanently attached to webpages, even if these ones may be dynamically modified). We currently see three trends emerging, each dwelling on the sociality of annotation.

First, web tools and applications like YouVersion (http://youversion.com) or BookGlutton (http://bookglutton.com) take a social networking approach to the study of texts and books. Individuals may select or buy texts and books from the website, invite friends or classmates to their online study group, and share their thoughts via comments, annotations and replies. Annotations can be made privately or publicly, allowing for private or group study. These tools focus on annotation as conversation. They emphasize and ease the creation of annotations online, and foster the analysis of the content by allowing the individual to leverage the community. In the case of BookGlutton, the interface is linked with a bookstore: the individual may only read and annotate books bought through that store. This may be an obstacle for students who want to upload different types of material to the site, or who may be looking for scholarly material not licensed for distribution through the bookstore.

The second trend concerns the annotation of any type of materials found on the web, from random webpages and blogs, to wikis and scholarly material. These tools, in turn, are not specific to a website. Rather, these are normally found as add-ons to different browsers. For instance, SideWiki (http://www.google.com/sidewiki), a Google project, is an add-on for Firefox and Chrome, allowing individuals to add notes to any types of websites. These annotations are public by default and displayed next to the webpage, in a sidebar viewable by anyone who has installed the widget. Here again, Google focuses on annotation as conversation. While this tool is not aimed at a scholarly public per se, SideWiki allows for annotations to be stored both on the original source and on the user’s Google profile, generating some type of handy notebook for the user.

The third trend directly involves scholarly communications: the development of online journal management and publishing systems. Projects such as the Open Journal System show great promise for academic publishing, but are still lacking some finesse in terms of reading tools. The
types of tools available to the users are left to the discretion of the publisher, but normally only involves linking articles to general bulletin boards where users can discuss broader issues.

While useful for the creation of annotation, these tools fail to integrate this activity within the overall lifecycle of annotation in the research context: they offer no means of purposively using these annotations in the research and writing context. While some of them support finer grained primitives, offering tools for comparing texts (YouVersion) and collating notes (SideWiki), most of these tools are however focused on annotation-as-reading, offering little help for further analysis and integration of content, not unlike desktop-based tools such as PDF readers. Furthermore, the first two trends are dependent on the availability of scholarly material both on the website itself (through the use of a bookstore) or on the Internet in general. This points to two specific requirements for the creation of digital tools supporting scholarship practices: the integration of the tool with scholarly publication and distribution networks, and the development of tools and features that not only support the creation of annotation, but that also foster the use and integration of these notes in the writing process of the scholar.

First, future tools supporting scholarship practices and assessment of source material must be integrated within the ecology of scholarly publications. This entails a close coupling of the tools with the publishing industry, mediated by the academic library. We can already see some institutions and services making general efforts towards this direction. Ontario Scholars Portal, a project created by the Ontario Consortium of University Libraries, offers shared services across the province, mainly purchasing and delivering electronic publications such as books and journals to students and faculty. Scholars Portal also offers an online reading platform, allowing students to obtain PDFs of journal articles and book chapters, as well as granting access to full-text HTML of these same resources via an interface tailored for online reading. Scholars Portal, as well as other vendors such as Ingram Digital, offers highlighting and note-taking tools for students and faculty members who register for this service.

However, there is also a need for these services to also be integrated with other scholarly activities such as bibliographic management. We should note the efforts of the Open Annotation Collaboration (http://openannotation.org/) project currently underway. This project, while focused on web annotation, attempts to link multiple information spaces such as vendors, bibliographic management software and annotation software, in order to integrate these and
create a hybrid space offering tools for annotation interoperable with these different services. This would let the scholar use the appropriate tools for her research without worrying about information fragmentation or loss of information and annotation when moving back and forth across services.

While upstream integration with vendors, libraries, and other services such as bibliographic software is crucial, future tools will also need to pay attention to the many activities involved in creating and using annotations and notes. Future digital environments for scholarly purposes should look at the finer grained activities that compose annotation and note-taking. These finer grained activities are truly the ones that are closer to the substance, to the materiality of the research object. For instance, as we have seen with our participants, using a note or an annotation does not strictly involve the simple assessment of the annotation. What we have seen is the necessity for participants to often not only assess the note, but to also refer to the original context, transfer the information to a new file, and finally to work in situ in the context of the master files. A digital environment supporting annotation must, therefore, recognize the heterogeneity and the necessity of these support activities, both when creating and using notes and annotations.
4 Conclusion

This study aimed to describe and analyze the annotation and note-taking practices of graduate students not only as material externalities of the research process, but also as crucial epistemic practices allowing students to progress from one research activity to the other. The goal of this research was threefold. First, we sought to understand how annotating and note-taking became epistemic practices in the context of the research project, addressing the duality of annotation as a relational dynamic as well as a partial material representation of the research object. Switching gears, we then examined annotations as “materialities of infrastructure”, revealing the different points of tension and contradiction in the practice, which brought to light certain aspects of the infrastructure supporting this epistemic practice. Finally, we situated annotation and note-taking in the context of the current digital textuality, pointing to areas of improvement for future information spaces supporting scholarship practices.

In order to fulfill these objectives, our research process led us to interview seven graduate students in humanities and social sciences faculties at a large Canadian university. The selected participants were all in the midst of a major individual research project, either completing a master’s thesis, submitting a major grant proposal, completing an independent study or finalizing a dissertation proposal. This ensured that participants created their own research process and allowed the researcher to assess representative events and incidents that were ultimately linked by the common thread of the project at hand. Interviews were supplemented by document collection. The researcher collected evidences of creation and use of annotations, from marked-up PDF files to outlines and drafts. These material representations of the research process allowed us to compare and contrast the student’s description of her research process with the actual physical and material externalities of this process.

We have furthermore designed a multi-perspectival framework, tailored to our research objectives and our definition of epistemic practice, in order to analyze the data collected from our interviews and documents. This framework allowed the researcher to zoom in and out of the annotation and note-taking practices by switching theoretical lenses. Consequently, we situated ourselves against all types of reductionism, revealing the complexity and messiness of the practice, without losing sight of our objectives or of our research questions. Zooming in, we observed and analyzed the doings of our participants, looking at the active role of material
elements and tools, the influence of the medium on the practice, as well as the methods and strategies performed by our participants. Knorr Cetina’s theory of epistemic practice and its focus on partial representations of the research object, as well as her understanding of the perpetual relational dynamic at play in such practice guided our efforts in this zoomed in view. Conversely, zooming out of the practice entailed the analysis of the wider repercussions of the practice on the overall scholarly activities and the research project. We traced the relationships between different practices, as well as revealed their structure by giving careful consideration to the social institutions bounding the practice. Theories of pragmatism and symbolic interactionism allowed us to look at the bigger picture and to trace relationships between the zoomed in and zoomed out perspectives. Our understanding of the zoomed out view and its relationship to the materialities of the practice was especially influenced by theories of the generalized others and of transactionality as discussed by Mead, and of ecologies of practice, social worlds and infrastructural inversions as put forth by Bowker and Star in various studies (Bowker, 1994; Bowker & Star, 1999; Star, 2002; Star & Bowker, 2006).

Zooming in, our results showed that students made use of a wide range of notes and annotations types, from meta notes revolving around organizational thoughts, to more interpretative notes, focused either on the reading at hand, or on the overall integration of material within the research project. These notes and annotations take on many forms and formats throughout the research project, and shift in quantity and importance as the student moves from one research activity to the other. Indeed, different activities call for different types of annotations. Therefore, we were able to discern between several types of reading activities taking place within the research process by using annotations and notes as units of analysis (see figure 47).

Figure 47: Spectrum of Annotation Types
We have also noticed an increase in structure and formalism of annotation content as the student progressed throughout the research project. These characteristics were in line with the thought development process of the student as one neared the completion of the research. Therefore, we noted that annotation and note-taking may not be necessarily only related to reading. Rather, these activities help the student orient herself towards the drafting and writing processes.

The student’s use of tools and software also paralleled this need for structure and formalism. Indeed, our participants used a wide assemblage of tools, favoring tools and software that allowed for freeform and easy note-taking in the beginning of the research process, slowly shifting to more structured tools towards the end of the project. An important finding was also that our participants’ practice was mostly performed digitally, using software and applications that allowed them to unify their materials and thoughts. However, this assemblage of tools and software was far from a streamlined system: our students used multiple workarounds, relied on several auxiliary tools (such as file management application) and even tailored generic tools for their scholarly needs (such as the browser). An important workaround strategy that our participants shared was the use of a text file to collate notes, ideas, quotes and bibliographic material from multiple sources, allowing the student to structure her thoughts on the fly and to unify her resources.

It is also important to note that paper-based activities were still present in our participants’ practices but to a lesser degree and for different purposes. Paper was mainly used to quickly offload one’s memory and played a transition role in the research process, allowing the student to move from one activity to the other by temporarily bridging both activities. Conversely, print-based annotations have also pointed to completely separate contexts of use and sets of relationships: students felt constrained to use paper-based annotations in a class or meeting setting, in the presence of individuals displaying authority over the student.

Zooming out, our study has revealed an interesting annotation lifecycle encompassing the creation and use of annotations (see figure 48). The activities composing the annotation lifecycle are tasks performing similar functions as Unsworth’s “scholarly primitives” (2000). This lifecycle can shift and be modified by the student according to the dominant scholarly activity at the time of creation or use of annotation. For instance, students who engaged in targeted reading
of source material towards the end of the research project tended to completely forego the Transfer, Maintain and Refer stages, instead opting for working their notes in situ.

Therefore, our research also allowed us to assess the overall structure of the research process, using annotation and note-taking practices as our units of analysis. Acquainting activities such as searching, retrieving, and reading revealed an intense period of creation of meta notes, which were almost immediately used on the spot. Conversely, research notes were predominant in analytical activities such as reading, organizing, interpreting, and outlining. Reading notes were generally created first during these activities. The student would then maintain these notes until a critical mass of information was gathered, supporting the critical thinking process of the student who then created more interpretive notes. The annotations and notes created during this phase also mirrored the increased structure and formalism in these analytical activities, from loose and unstructured reading notes to more formal outlines collating the student’s ideas as the project marched on. Finally, composition activities such as drafting and writing were characterized by the use of research notes, the creation and almost immediate use of meta notes and the shortening of the lifecycle by foregoing Transfer, Maintain and Refer activities in order to work material in situ.

![Annotation Lifecycle Diagram](image-url)
Interpreting these results using our multi-perspectival framework allowed us to fulfill our three objectives. First, looking at annotation and note-taking as epistemic practices allowed us to reveal the function of the “micro-practice” or the “primitive practice”. Indeed, our study demonstrated how the processes of annotating and note-taking are situated between different levels of activities and granularities. Using Unsworth and Palmer & Cragin’s concept of scholarly or information work primitives, we discerned several finer-grained activities composing annotation and note-taking practices. These activities, the proper “scholarly primitives” are simple tasks that need to be combined in order to create or use an annotation. Conversely, we noted the integration of annotation and note-taking activities within larger realms, acting as support activities for reading, analyzing and writing. We have therefore termed annotation and note-taking as “primitive practices”. These primitive practices have several epistemic qualities. We highlighted the bridging function of annotation, the relational dynamic that pushes the student forward in the research project by mediating not only between the research object and the material representations of the object, but also between different activities and tasks throughout the research project. Interestingly, this dynamic creates externalities, the annotations themselves, which are partial representations of the never-complete research object. The lack defining these representations points the researcher towards the missing pieces of the puzzle, always re-creating the relational dynamic of the epistemic practice.

Our study also revealed several points of tensions and negotiations during this epistemic practice that hint at the close encounter of the individual with the infrastructure supporting the practice. These points of tension perpetuate, maintain and slightly shift the infrastructure of the practice. Specifically, we have noted how current digital tools claimed to be designed for fostering an active reading strategy indeed hindered this process and perpetuated a traditional notion of annotation-as-reading. The structure and formalism of these tools points to the passive assumptions built into the software and applications used by the students. This also explains why our participants deployed so many workaround techniques in order to sustain an active reading strategy geared towards analytical and writing activities. Our research further shed light on the various criteria at play when selecting the appropriate medium and tools for scholarly activities. Students transacted and negotiated with themselves when selecting the appropriate tool, often favoring the speed and efficiency of annotation over structure and formalism in the beginning of the project and vice versa during drafting activities. Moreover, the choice of maintaining a
strictly digital practice was influenced by the student’s need to unify all resources, as well as the overall chain of institutions favoring digital services. Finally, social and disciplinary norms also shaped the annotation practice: students, advisors and institutions were members of several intersecting social worlds and played different roles in each of them, often imposing different expectations on the student.

Finally, situating annotation and note-taking practices in the recent rise of digital textuality allowed us to take a closer look at the current services supporting screen-based annotation. We noticed a severe lack of tools dedicated to scholarly work online. The current tools supporting annotation and note-taking dwell on the traditional annotating-as-reading model, making weak efforts to support finer-grained activities involved in annotating and wider processes, especially analytical and composition activities. We furthermore noted that most of these services are dependent upon publishers and vendors supplying scholarly resources to the website or web application. Hence, we proposed specific requirements for the creation of future digital tools supporting scholarship practices: the integration of the tools within the scholarly publication ecology, and the development of tools and features that not only support the creation of annotation, but that also foster the use and integration of these notes in the writing process of the student.

4.1 Limitations of the Study

The multi-perspectival approach used in this research allowed the researcher to elicit interesting findings and implications unaddressed by previous studies of annotation and note-taking. Indeed, zooming in and out of this epistemic practice framed this beloved academic topic in a new way, distancing ourselves from traditional studies of annotation-as-reading by looking not only at contexts of creation of annotation, but also at contexts of use. Hence, we revealed that annotation and note-taking were processes helping the student to orient herself towards drafting and writing activities, bridging the gaps between different scholarly activities, and reconciling partial representations of the research object with the student’s conception of the research object itself. There were, however, some limitations to our research that need to be addressed.

As discussed in our methodology, the inherent limitations of the small sample size and the interview technique used, may prevent us from generalizing our findings. Our results however allow for new lines of questioning to emerge, specifically around the scholarly practices of the
new generation of graduate students and scholars who grew up with new technology. While our research design involving interviews and document collection did set us apart from a tradition of studies of annotation focused either on usability studies or on a strict examination of books and textbooks, we might have been able to elicit more complete findings by shadowing our participants throughout their daily activities, for an extended period of time. This might have been especially useful for further dissecting the finer-grained activities involved in annotating and note-taking, and would have allowed us to gain a greater understanding of the working of annotations in situ, in the context of writing activities. Our current study only hints at the great wealth of information activities and strategies involved in drafting, versioning and writing.

Furthermore, and as a consequence of our project’s scope, our study is restrained to individual research projects, only superficially addressing the relationships between this practice and annotation activities in other academic contexts such as group work and class work. This also points to another limitation: the assessment of the research paper as a genre. Indeed, all of our participants were working on individual research papers. Perhaps the annotations and note-taking practices might have been slightly different if students were working towards a conference presentation, a conference poster or a major grant proposal.

Finally, it appears that our project could have been divided in several studies. While our current research yields a global perspective on annotation, revealing interesting areas of the practice that are not commonly assessed, such as annotation-as-writing, our groundwork covered multiple facets that might have been better assessed individually. For instance, investigating the creation and use of annotation could have been two separate studies altogether, and perhaps the examination of the annotation lifecycle and its response to various scholarly activities would be productive in itself. Because annotation and note-taking are intertwined with many scholarship practices, it was quite challenging to address all facets and implications in thorough manner over the course of one research project.

4.2 Future Work

The limitations to our current research also serve as recommendations for future research. As discussed, focusing on annotation-as-writing and the subtending activities involved in this epistemic practice may be fruitful for the field of information in general. Indeed, while many studies have dedicated time and efforts to reveal the intricacies of scholarly reading and the
consequences of new technologies on the habits of students, very few researchers have addressed the role of writing activities in the research project, preferring to avoid this grey zone.

This current study could also be used as a springboard for future research on digital tools sustaining annotation and note-taking practices in a research context. Our research revealed a gap in the availability and usefulness of tools sustaining scholarship practices dealing with online resources. Hence, new tools and interfaces should pay attention to the use and integration of annotations in the writing process of students and academics, as well as potentially support an hybrid practice involving resources from different provenance, from web-only material and downloadable files, to printed resources. This may also involve looking at how scholarship practices may be performed using new devices such as e-book readers, tablet computers and smartphones, and in what capacities these new platforms may sustain scholarly activities. Integrating these devices within the ecology of scholarly work will be crucial in order to minimize information fragmentation.

Finally, future work may also involve looking at the annotation and note-taking practices of other social groups such as well-established academics, students in natural sciences or even professional knowledge workers. We believe that epistemic practices may have slightly different characteristics according to the group under study and thus annotation may play a very different role compared to the bridging function as we revealed in this study.
References


Dillon, A. (2010, June). *As We May Have Thought, and May (Still) Think*. Keynote Address presented as the ACM Hypertext 2010 Conference, University of Toronto, Toronto, Canada.


Appendices

Appendix A: Recruitment Letter for Interviews and Document Collection

Hello *(insert name of contact)*,

I am a Master student at the University of Toronto working under the supervision of Dr. Matt Ratto at the Faculty of Information. *(If reference is from personal network … *(insert name of referee)* was kind enough to give me your name and e-mail address. *(If reference is from faculty and staff directory… I am taking the liberty to contact you because of your current position as *(insert position)* at the *(insert name of the faculty or organization)*.)*

I would like to invite you to participate in a study I wish to do for my thesis. The aim of my thesis is to understand how the rise of digital media has shaped and modified the way the academic community engages with publications. My research thus focuses on how different members of the academic publishing community engage with texts both intellectually and physically (e.g. via highlighting, note-taking, etc…) as part of their professional activities.

As a participant, you will be interviewed twice on the site where you perform most of your daily professional activities, for sessions of approximately one hour. In the course of these sessions, documents such as highlighted texts or notebooks may be collected or photographed in order to assess the physical dimension of intellectual work. You will also be asked to participate in a group workshop of approximately two hours with other members of the academic publishing community in order to reflect on certain topics related to the current state of scholarly publications. The interviews and workshop are currently scheduled to take place between the month of November and January, at your convenience. The results of this study will be of great value to the academic publishing community, providing a new perspective on the current use of academic journals and insights for future publishing models and platforms.

I am well aware that some of the documents or conversations may contain sensitive or confidential information. Any confidential or identifying information will be anonymized and only the researcher and her supervisor will have access to the original data collected. I would like to assure you that the study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Toronto.

If you are interested in participating in this research, or would like to further discuss my study or this letter, please contact me at me.belanger@utoronto.ca, or by phone at (647) 520-0675. Alternatively, you may contact my supervisor, Dr. Matt Ratto, at matt.ratto@utoronto.ca.

Thank you for your consideration,

Marie-Eve Bélanger,
Master student, Faculty of Information, University of Toronto
Appendix B: Interview Guide

1. Professional activities / tasks related to texts and journal articles
   • What are the key aspects of your work?
   • What are the tasks you normally accomplish that are linked to the text/journal article itself?
   • Can you show me or describe those tasks?

2. Medium used to accomplish tasks
   • How often do you use the computer to accomplish the tasks you described?
   • Do you need to print out the text or the journal article in order to accomplish any tasks?
   • Why would you print out a text or a journal article?
   • What can you do with printouts that you cannot do on screen? And conversely, what can you do on screen that you cannot do with printouts?
   • Have you used other systems (digital or analog) to accomplish these tasks?
     o If so, what type of system was it?
     o Was it imposed by your employer, someone else, or did you implement it yourself, to accommodate your personal routine?

3. Textual Practices (Micro level, independent of the medium of choice)
   • How do you bring attention to specific parts of the text?
   • What are the different types of annotations you generally make on a text?
   • Do you use markers, highlighters, pens, pencils or any other tools when working with a text on paper?
   • Do you use specific programs and applications when working on a text on screen?
   • What type of marks would you normally make on a text on paper? What type of marks would you normally make on a text on screen?
   • Are these types of marks different, physically and meaning-wise, whether you’re working on screen or on paper?
   • Is color important for you when working with a text?
     o If so, can you explain the meaning of the ink colors you are using?
   • Do you store notes or other marks elsewhere than on the text?
   • When marking and annotating on paper or on screen directly on a text, how do you control the amount of annotations you make? When are annotations distracting you from the text?
   • If marking and annotating on paper, do you normally also mark and annotate the digital version of the text? (and vice versa)
     o If so, are the annotations similar?

4. Textual Practices (Intermediate level, independent of the medium of choice)
   • If you had to switch activity or be distracted for a few instants from the text you are working on, how would you recall where you were in the text?
   • Do you use bookmarks or any other signaling devices to mark your place in a text?
   • Do you use multiple methods?
     o If so, how do they differ physically and meaning-wise?
   • What would count as your preferred marker in a text? A page number, the beginning of a paragraph, a precise sentence on a page, or anything else?
   • Once you recall your marker, where do you generally start reading? At the marker itself? A few pages earlier? The beginning of the article?
   • If you bookmark a page on an article printout, would you translate this bookmark on the digital version also?

5. Textual Practices (Macro level, independent of the medium of choice)
   • As part of your daily activities, do you handle multiple texts or journal articles at once?
   • How do you manage multiple texts at once?
• Do you have a personal system that allows you to keep track of the texts you are currently working with?
  o If so, can you explain and describe this system?
• When dealing with multiple texts at once, is there generally a link (conceptual, thematic, chronological, etc.) between the texts?
• What happens to the text or the journal article, printed or digital, once you are done with it?
• If you have both a digital and printed version, do you keep both once you are done with it?
• If you need to re-access a text some time after your first contact, how do you recall where the text is stored or what the text was about?
  o Do you have a specific classification system that you are currently using?
  o Is this system personal? Institutional?

6. Professional activities related to other stakeholders
• Who do you interact with as part of your daily activities?
• Who is necessary to the full completion of your daily tasks?
• When working on a text or a specific journal article, do you normally collaborate with other people to accomplish your task?
  o How would you describe these people? Are they colleagues, manager, employees?
  o Are they mostly from within your organizations?
• In order to complete your daily task, do you need to interact with people from outside of your organization?
  o How would you describe those people?
  o Are they colleagues, managers, employees, suppliers, contractors?
  o What type of relationship do you have with the people you mentioned?

7. Communication and power structure of the community
• Where do you situate your position in the chain of (production/distribution) of journal articles?
• Who is your manager or your superior?
• Who briefs you on tasks that need to be accomplished?
• Do you manage or brief other employees on tasks that they need to accomplish?
• How do you communicate with other employees regarding work tasks?
  o Is this system imposed by your employer?
  o What would be your preferred way of communicating with other employees?

8. Collaboration and textual practices
• Do you collaborate with colleagues to accomplish specific tasks?
  o If so, which task do you accomplish collaboratively?
  o What is the professional title of your other colleagues?
• Do they perform a similar task as you on the text or something different?
• How do you manage who controls the document?
• Do you use a specific versioning system or protocol?
  o Is this system or protocol based on a personal preference or is it imposed by your organization?
• How do you manage the comments, marks or annotations of multiple users on a single document?
• How do you differentiate the comments, marks and annotations of multiple users on a single document?
• What software or application do you use to perform tasks collaboratively?
• What features do you use if you work collaboratively with someone?
Appendix C: Ethics Approval

### University of Toronto
**Office of the Vice-President, Research**
**Office of Research Ethics**

**PROTOCOL REFERENCE # 24622**

November 20, 2009

Dr. Matt Ratto
Faculty of Faculty of Information
University of Toronto
45 Willcocks St.
Toronto, ON M5S 1C7

Ms. Marie-Eve Belanger
Faculty of Information
University of Toronto
140 St. George St.
Toronto, ON M5S 3G6

Dear Dr. Ratto and Ms. Belanger:

Re: Your research protocol entitled “The Evolving Ecology of the Academic Publication: Understanding the infrastructure of emergent academic textual practices and their implications for future information spaces”

<table>
<thead>
<tr>
<th>ETHICS APPROVAL</th>
<th>Original Approval Date: November 20, 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expiry Date: November 19, 2010</td>
</tr>
<tr>
<td></td>
<td>Continuing Review Level: 1</td>
</tr>
</tbody>
</table>

We are writing to advise you that a member of the Social Sciences, Humanities & Education Research Ethics Board has granted approval to the above-named research study, for a period of one year, under the REB’s delegated review process. **Please ensure that you submit an Annual Renewal Form or a Study Completion Report 15 to 30 days prior to the expiry date of your study. Note that annual renewals for studies cannot be accepted more than 30 days prior to the date of expiry, as per federal and international policies.**

All your most recently submitted documents have been approved for use in this study.

**Any changes to the approved protocol or consent materials must be reviewed and approved through the amendment process prior to its implementation. Any adverse or unanticipated events should be reported to the Office of Research Ethics as soon as possible.**

If your research has funding attached, please contact the relevant Research Funding Officer in Research Services to ensure that your funds are released.

Best wishes for the successful completion of your project.

Yours sincerely,

Daniel Gyewu
Research Ethics Coordinator
Appendix D: Letter of Consent sent to Participants

Research project title: The Evolving Ecology of the Academic Publication

Research team:
Principal investigator: Marie-Eve Bélanger
Master student, Faculty of Information, University of Toronto
647-520-0675
me.belanger@utoronto.ca

Supervisor: Dr. Matt Ratto
Assistant Professor, Faculty of Information, University of Toronto
416-946-5415
matt.ratto@utoronto.ca

Summary of the research project
Increasingly, journal articles and other scholarly materials are available to scholars in digital form. The rise of digital media in the context of the academic publication has yielded a compendium of new means of production, distribution and consumption of texts. This new environment can be currently understood as a type of playground where various actors such as readers, authors, editors and publishers are experimenting with the affordances of the medium. While some of these actors primarily read and enact their professional practice via computer screens, others find it necessary to print out digital documents. One reason for this printing may be the need to highlight, annotate, and mark-up materials as part of their professional activities. My research thus aims to understand how readers, authors, editors, librarians, designers, proofreaders and publishers engage with texts intellectually and physically.

The objective of my study is to assess how the current structure of the academic publishing community affects the types of engagements one can have with a text in the course of their professional activities. Thus, the interviews and workshop planned for this study will help me to understand the wide variety of intellectual and physical engagements with text and to examine how the structure of the community and relationships between different stakeholder affect these engagements.

Invitation to participate and respect of ethical principles
As a stakeholder involved in the academic publishing community, you are invited to participate in this study. (For participants invited to interviews and ‘critical making’ session) Part of this participation consists of being interviewed twice on the site where you perform most of your daily professional activities, for sessions of approximately one hour. In the course of these sessions, documents such as highlighted texts or notebooks may be collected or photographed in order to assess the physical dimension of intellectual work. You will also be asked to participate in a group workshop of approximately two hours with other members of the academic publishing community in order to reflect on certain topics related to the current state of scholarly publications. (For participants invited to a ‘critical making’ session only) You will be asked to participate in a group workshop of approximately two hours with other members of the academic publishing community in order to reflect on certain topics related to the current state of scholarly publications.
Please be assured that all collected information will be treated confidentially. Thus, people who could have access to this information, namely Marie-Eve Bélanger and Matt Ratto, have signed a confidentiality agreement. Transcripts from the individual interviews and ‘critical making’ session, as well as field notes will be kept in a secure location for at most five years and then destroyed.

While all collected data, including verbal interactions and written documents may be used for publication purposes, the anonymity of (name of the organization) and individual participants will be ensured outside of your organization by designating stakeholders by surrogate names in interview and ‘critical making’ session transcripts, field notes, and any other collected data. Considering the focus of the study and the confidentiality measures that will be taken, participating in this pilot study should not cause you any prejudice. While this study will give you the opportunity to interact with other stakeholders and envision future publishing models and platforms, it should not bring you any direct benefit either. However, although your organization has agreed to the conduct of this study, you should not, under any circumstance, feel obliged to participate; I assure you that your decision will not be communicated to any of your colleagues or superiors. Moreover, once you have agreed to participate, you may withdraw at any time without further justification. Your participation should be completely voluntary.

Please note that you can accept to participate solely to the individual interviews, or to a ‘critical making’ workshop; you can also agree to participate to both activities. You may also decline to answer any question during the interviews or ‘critical making’ session. Finally, although it is planned to do audio recordings of individual interviews and video recordings of the ‘critical making’ workshop, this will only be done if all participants agree to it at the time of the activity.

If you want further information on the study, you may contact the main investigator or her supervisor. If you want more information regarding your rights as a participant, you may contact the Office of Research Ethics by phone at 416-946-3273 or by email at ethics.review@utoronto.ca. You may also keep a copy of this letter for your records.

**Participant signature**

Having read and understood the above text, and having had the possibility to ask and receive complementary information on the study, I consent to participate to the following activities in this research:

- Individual interview: ____________________
- ‘Critical Making’ workshop: ____________________

I would like to receive a copy of the final thesis document and a copy of any published materials using results from the study:

- Yes _____________       No ______________

Participant name: ____________________
Participant signature: ____________________
Date: ____________________
### Appendix E: Table summarizing types of notes and research activities

<table>
<thead>
<tr>
<th>Activities</th>
<th>Types of Notes</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideation and early thought development</td>
<td>Meta Notes, mainly project-related</td>
<td>Very temporary</td>
</tr>
<tr>
<td></td>
<td>Ideas and thoughts (mindmaps)</td>
<td>May be kept throughout the research project and act as a guide for the student, may shift in content and become increasingly formalized (e.g. from mindmap to outline)</td>
</tr>
<tr>
<td>Skimming</td>
<td>Selection (highlight)</td>
<td>Highlights tends to be created both for attentional purposes and to flag potentially interesting passages.</td>
</tr>
<tr>
<td></td>
<td>Comments (summaries)</td>
<td>Content tends to be a general appreciation of the article.</td>
</tr>
<tr>
<td>Squirreling &amp; Bouncing</td>
<td>Selection (copy and paste)</td>
<td>Default technique since proper annotation and commenting is hindered by the browser technology</td>
</tr>
<tr>
<td></td>
<td>Comments (summaries)</td>
<td>In external document. Content tends to be a general appreciation of the article. Derived from the skimming activity happening concurrently.</td>
</tr>
<tr>
<td>Targeted Reading</td>
<td>Selection (highlight)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Selection (copy and paste)</td>
<td></td>
</tr>
<tr>
<td>Reading for class</td>
<td>Selection (highlight)</td>
<td>Formal qualities may be different due to the constraints of the class setting; content in line with class topics</td>
</tr>
<tr>
<td></td>
<td>Comments (questions &amp; marginalia)</td>
<td>Formal qualities may be different due to the constraints of the class setting; content in line with class topics</td>
</tr>
<tr>
<td>Background reading and reading around</td>
<td>Selection (highlight)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Selection (copy and paste)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comments (questions &amp; marginalia)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ideas and thoughts</td>
<td></td>
</tr>
<tr>
<td>Reading for deep comprehension</td>
<td>Selection (highlight)</td>
<td>Self-limitation to one annotating strategy that is deemed non-intrusive; Textual comments perceived as a distraction</td>
</tr>
<tr>
<td>Activities</td>
<td>Types of Notes</td>
<td>Characteristics</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Conducting fieldwork</td>
<td>Comments</td>
<td>May take the form of “tagging” if the student is using a qualitative research software</td>
</tr>
<tr>
<td>Analyzing</td>
<td>Ideas and thoughts</td>
<td>More formal notes, based on the accumulated knowledge base. Creation of outlines</td>
</tr>
<tr>
<td></td>
<td>Meta notes, mainly content-related</td>
<td>Notes pointing out gaps in the research and acting as personal reminders</td>
</tr>
<tr>
<td>Writing and dissemination</td>
<td>Comments</td>
<td>Often formalized and structured for possible dissemination to other individuals</td>
</tr>
<tr>
<td></td>
<td>Selection (copy and paste)</td>
<td>Often integrated directly into the draft</td>
</tr>
<tr>
<td></td>
<td>Meta Notes</td>
<td>Proofreading and editing notes, identification of gaps in the content. These notes are appended to the material currently being written by the student (instead of the original source material)</td>
</tr>
</tbody>
</table>