



# **Integrating Printed and Digital Documents**

**Interaction Models and Techniques for Collaborative Knowledge Work**

## **Dissertation**

Zur Erlangung des akademischen Grades  
eines Doktor rerum naturalium (Dr. rer. nat.)

Eingereicht von  
**Jürgen Steimle**  
geb. in Mutlangen

Angenommen vom Fachbereich Informatik  
der Technischen Universität Darmstadt

Gutachter: Prof. Dr. Max Mühlhäuser (TU Darmstadt)  
Prof. Dr. Jan Borchers (RWTH Aachen)  
Prof. Dr. Werner Sesink (TU Darmstadt)

Tag der Einreichung: 18.05.2009  
Tag der Disputation: 23.06.2009

Darmstadt 2009  
D17

Please cite this document as follows:

URN: urn:nbn:de:tuda-tuprints-19233

URL: <http://tuprints.ulb.tu-darmstadt.de/1923>

This document is provided by tuprints,  
the e-publishing service of TU Darmstadt.  
<http://tuprints.ulb.tu-darmstadt.de>  
[tuprints@ulb.tu-darmstadt.de](mailto:tuprints@ulb.tu-darmstadt.de)

## Abstract

Today's knowledge work is characterized by the simultaneous use of paper and digital documents. This implies disruptive transitions from paper to digital media and vice versa, particularly because interaction techniques differ in the two worlds. For instance, traditional references are created and followed differently than digital hyperlinks. This thesis advances the integration of paper and digital documents. It contributes empirically-grounded conceptual work in the fields of interaction theory, interaction techniques and interactive systems.

Three *initial field studies* explore how paper and digital media are used in learning at universities. They provide evidence for the high importance of paper and show that, in many settings, the use of printed information exists in parallel with using digital documents. Based on these results, an *ecological perspective* of knowledge work is chosen as the approach for developing the theoretical basis of this thesis. This perspective advocates an integral view on the ensemble of collaborating users, of physical and digital artifacts, of work practices and of their interplay.

By generalizing the findings of the empirical field studies and based on the ecological perspective, we develop a *theoretical interaction model of Pen-and-Paper User Interfaces (PPUIs)*. This model is of analytical value and provides guidance on how to design PPUIs that are easy and intuitive to use. Its underlying principle is a systematic separation between a semantic and a syntactic level of interaction. Based on this separation, syntactic interaction primitives are identified, which act as building blocks for PPUIs that support semantic activities.

Furthermore, we contribute *novel interaction techniques and visualizations* for cross-media knowledge work with documents. These are based on the interaction model and provide support for annotating, linking and tagging, all in a hybrid setting of printed and digital documents. First, in the field of *paper-based annotation*, we introduce the concept of user-adaptable printed interfaces. Moreover, we significantly advance asynchronous sharing by presenting a paper-based mechanism for sharing annotations and a visualization that integrates handwritten annotations of multiple users. Second, we contribute a pen-based interaction technique for creating and following *cross-media hyperlinks* between printed and digital documents. The same digital pen and the same interactions can be used both on printed documents and on digital documents on a tabletop screen. Third, we contribute four novel techniques for *tagging documents and processes*. These offer a rich user experience being inspired by the varied practices of using paper. They leverage tangible stickers, paper cards for defining and applying tags, printed buttons, and other tangible objects.

Moreover, this thesis contributes CoScribe, a *system framework for collaborative cross-media knowledge work*. This integrates the interaction techniques and visualizations into a consistent and coherent concept. Following the ecological perspective, CoScribe covers entire workflows with a strong emphasis on *collaboration*.

Finally, we present *evaluation results*. We implemented a *working prototype* of CoScribe, which was used in three evaluation studies. Their results provide evidence that CoScribe significantly enhances both work performance and user satisfaction.

## Zusammenfassung

Trotz großer Fortschritte bei digitalen Medien und derer Vorteile wird traditionelles Papier in der Wissensarbeit wegen seiner besonderen Charakteristika weiterhin in großem Umfang verwendet. Die gleichzeitige Verwendung beider Dokumentenarten führt zu bruchhaften Übergängen. Beispielsweise unterscheiden sich traditionelle Referenzen von digitalen Hyperlinks. Diese Dissertation hat zum Ziel, die Kluft zwischen Papier und digitalen Dokumenten zu verringern. Sie beinhaltet Beiträge in Interaktionstheorie, Interaktionstechniken und interaktiven Systemen.

Drei *einleitende Feldstudien* untersuchen, wie Papier und digitale Medien im universitären Lernen verwendet werden. Sie zeigen die hohe Bedeutung von Papier und ergeben, dass in vielen Situationen gedruckte parallel mit digitalen Dokumenten verwendet werden. Anhand dieser Ergebnisse wird eine *ökologische Perspektive* auf Wissensarbeit als Ansatz für die theoretische Basis dieser Arbeit gewählt. Diese vertritt eine integrierte Sicht auf das Ensemble von kooperierenden Nutzern, physischen und digitalen Artefakten, Arbeitspraktiken und deren Zusammenhänge.

Als ein weiterer Beitrag wurde ein *theoretisches Interaktionsmodell für Stift- und Papier Nutzerschnittstellen (PPUIs)* entwickelt. Dieses basiert auf einer Generalisierung der Ergebnisse der Feldstudien und auf der ökologischen Perspektive. Es unterstützt sowohl Analyse als auch Entwurf von PPUIs. Dem Modell liegt eine systematische Trennung zwischen einer semantischen und einer syntaktischen Interaktionsebene zu Grunde. Auf dieser Basis werden syntaktische Interaktionsprimitive identifiziert. Diese dienen als Bausteine für PPUIs, die spezifische semantische Aktivitäten unterstützen.

Ein wesentlicher Beitrag dieser Dissertation sind *neue Interaktionstechniken und Visualisierungen* für Wissensarbeit, die gedruckte und digitale Medien integrieren. Diese basieren auf dem Interaktionsmodell und erlauben es, Dokumente zu annotieren, zu verlinken und mit Tags zu versehen. Im Bereich der *papierbasierten Annotation* führen wir nutzeranpassbare gedruckte Schnittstellen ein. Außerdem wird die asynchrone Kooperation deutlich verbessert durch einen papierbasierten Mechanismus für den Austausch von Annotationen mit anderen Nutzern und durch eine Visualisierung, die handschriftliche Annotationen vieler Nutzer integriert darstellt. Zweitens wird eine stiftbasierte Interaktionstechnik für *Medien-überspannende Hyperlinks*, die gedruckte und digitale Dokumente integrieren, vorgestellt. Dabei kann derselbe digitale Stift sowohl auf Papier als auch auf einem Tabletop-Bildschirm verwendet werden. Drittens werden *neue Interaktionstechniken für das Tagging von Dokumenten und Prozessen* eingeführt. Inspiriert durch die variierten Praktiken Papier zu verwenden, bieten diese eine intuitive und vielfältige Interaktion. Sie basieren auf physischen Aufklebern, Papierkarten, gedruckten Schaltflächen sowie weiteren physischen Objekten.

Des Weiteren führt diese Dissertation CoScribe ein, ein System-Rahmenwerk für kollaborative Wissensarbeit in hybriden Dokumentenräumen. CoScribe integriert die Interaktionstechniken und Visualisierungen in ein konsistentes und kohärentes Konzept. Basierend auf der ökologischen Perspektive deckt CoScribe komplette Arbeitsabläufe ab und bietet Unterstützung für Kooperation mehrerer Nutzer.

Abschließend stellen wir *Evaluationsergebnisse* vor. Dafür wurde ein *Prototyp des CoScribe-Systems* implementiert und in drei Nutzerstudien getestet. Die Ergebnisse zeigen, dass CoScribe die Effizienz und Nutzerzufriedenheit signifikant erhöht.

# Acknowledgments

This work would not have been possible without the support and encouragement of my advisor, colleagues and students. I am grateful to all those who contributed to the outcome of this thesis.

First and foremost, I would like to thank Max Mühlhäuser (TU Darmstadt), my advisor, for his unlimited support and excellent advice. I am also grateful to Jan Borchers (RWTH Aachen) and to Werner Sesink (TU Darmstadt) for fruitful discussions and for acting as co-referees. I wish to extend special thanks to Oliver Brdiczka (Palo Alto Research Center) for his continuous advice.

I am grateful for having been a member of the interdisciplinary postgraduate school “eLearning” at TU Darmstadt, the context in which my work was funded. I thank all advisors and all Ph.D. students of the postgraduate school for the highly inspiring atmosphere and the supportive discussions. In particular, I would like to thank Christoph Koenig for his cross-disciplinary advice.

Many thanks are due to all present and former members of the Telecooperation group at TU Darmstadt. They provided a very friendly place to work and supported this thesis in many respects. I also would like to thank the students who contributed to the implementation of the prototype system. These were Stefan Buhrmester, Roman Lisserman, Simon Olberding, Michael Stieler, Sasa Vukancic and Jie Zhou.

I gratefully acknowledge Scott Klemmer, Andreas Paepcke and Ron Yeh (Stanford University) for their advice and feedback on my work during my lab visit.

Martin Donner (Humboldt University of Berlin) supported this work by designing the audio feedback.

Last but not least, I am highly grateful to my parents for their love and support.



# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>1</b>  |
| 1.1      | Motivation . . . . .  | 1         |
| 1.1.1    | The Gap between Paper and Digital Information . . . . .                           | 1         |
| 1.1.2    | Knowledge Work with Documents . . . . .   | 2         |
| 1.2      | Contributions of this Thesis . . . . .  | 4         |
| 1.3      | Research Method . . . . .   | 6         |
| 1.4      | Publications . . . . .  | 8         |
| 1.5      | Structure of this Thesis . . . . .  | 9         |
| <b>2</b> | <b>Knowledge Work with Paper Documents</b>  | <b>13</b> |
| 2.1      | Scenario: Learning and Knowledge Work at Universities . . . . .                   | 14        |
| 2.2      | Affordances of Paper Documents . . . . .  | 15        |
| 2.3      | Field Studies . . . . .   | 19        |
| 2.3.1    | Study I: Notetaking in Lectures . . . . .   | 19        |
| 2.3.2    | Study II: Annotating with Pen and Paper vs. Annotating with<br>a Laptop . . . . . | 28        |
| 2.3.3    | Study III: Media Use and Collaboration in Learning Groups . . . . .               | 34        |
| 2.3.4    | Summary and Implications . . . . .  | 37        |
| 2.4      | An Ecological Perspective of Knowledge Work . . . . .                             | 38        |
| 2.4.1    | Distributed Cognition . . . . .   | 39        |
| 2.4.2    | Information Ecologies . . . . .   | 40        |
| 2.4.3    | Ecological Perspective . . . . .  | 41        |
| 2.5      | Requirements . . . . .  | 42        |
| 2.6      | Conclusions . . . . .   | 47        |
| <b>3</b> | <b>State of the Art</b>   | <b>51</b> |
| 3.1      | Theoretical Models of Pen-and-Paper User Interfaces . . . . .                     | 52        |
| 3.1.1    | Models of Tangible User Interfaces . . . . .                                      | 52        |
| 3.1.2    | Models of Pen-and-Paper User Interfaces . . . . .                                 | 54        |
| 3.1.3    | Discussion . . . . .  | 54        |

|          |   |            |
|----------|---|------------|
| 3.2      | Technology . . . . .  | 55         |
| 3.2.1    | Page Identification and Location Tracking . . . . .                         | 55         |
| 3.2.2    | Digitizing the Contents of Paper Documents . . . . .                        | 57         |
| 3.2.3    | Tracking User Interaction on Paper . . . . .                                | 58         |
| 3.3      | Pen-and-Paper Applications . . . . .  | 64         |
| 3.3.1    | Paper-based Notetaking . . . . .  | 64         |
| 3.3.2    | Paper-based Annotation . . . . .  | 67         |
| 3.3.3    | Paper-based Hyperlinking . . . . .  | 69         |
| 3.3.4    | Paper-based Tagging . . . . .   | 76         |
| 3.4      | Discussion and Conclusions . . . . .  | 79         |
| 3.4.1    | Interaction Paradigms and Metaphors for the Conceptual Activities . . . . . | 80         |
| 3.4.2    | Perceptual Coupling of Printed and Digital Information . . . . .            | 82         |
| 3.4.3    | Collaboration . . . . .   | 83         |
| 3.4.4    | Integration into Existing Information Ecologies . . . . .                   | 84         |
| <b>4</b> | <b>Interaction Model of Pen-and-Paper User Interfaces</b>                   | <b>87</b>  |
| 4.1      | Pen-and-Paper User Interfaces . . . . .                                     | 88         |
| 4.2      | Interactions . . . . .  | 90         |
| 4.2.1    | Semantic and Syntactic Levels of Interaction . . . . .                      | 90         |
| 4.2.2    | Semantics: Conceptual Activities . . . . .                                  | 92         |
| 4.2.3    | Syntax: Core Interactions . . . . .   | 92         |
| 4.2.4    | Applicability of Core Interactions to Related Work . . . . .                | 95         |
| 4.2.5    | Mapping Between Syntax and Semantics . . . . .                              | 96         |
| 4.3      | Information . . . . .   | 99         |
| 4.3.1    | Document Types and Representations . . . . .                                | 99         |
| 4.3.2    | Relations between Physical and Digital Information . . . . .                | 100        |
| 4.3.3    | Appropriate Complementing Representations . . . . .                         | 103        |
| 4.4      | Collaboration . . . . .   | 105        |
| 4.4.1    | Local Collaboration . . . . .   | 105        |
| 4.4.2    | Remote Collaboration . . . . .  | 107        |
| 4.4.3    | Privacy . . . . .   | 108        |
| 4.5      | Conclusions and Design Guidelines . . . . .                                 | 108        |
| <b>5</b> | <b>CoScribe: Collaborative Cross-media Annotation, Linking and Tagging</b>  | <b>111</b> |
| 5.1      | Overview of CoScribe . . . . .  | 113        |
| 5.1.1    | Pen-based Interaction with Printed and Digital Documents . . . . .          | 113        |
| 5.1.2    | A Toolset for Document-based Knowledge Work . . . . .                       | 114        |
| 5.1.3    | Co-located and Remote Collaboration . . . . .                               | 116        |
| 5.2      | Integrating Printed and Digital Documents . . . . .                         | 118        |
| 5.2.1    | A Unified Interaction Device . . . . .                                      | 118        |
| 5.2.2    | Mobile Use and Real-Time Interaction . . . . .                              | 121        |
| 5.2.3    | Printed Documents and Tools . . . . .                                       | 121        |
| 5.2.4    | Digital Document Viewer . . . . .   | 123        |

---

|          |   |            |
|----------|---|------------|
| 5.2.5    | Ecological View . . . . .   | 124        |
| 5.2.6    | Transformations between Printed and Digital Representations . . . . . | 127        |
| 5.3      | Annotating Documents . . . . .  | 130        |
| 5.3.1    | Handwritten Free-Form Annotations . . . . .                           | 132        |
| 5.3.2    | Adaptivity of the Printed User Interface . . . . .                    | 134        |
| 5.3.3    | Paper-based Sharing of Annotations . . . . .                          | 135        |
| 5.3.4    | Visualization of Shared Annotations . . . . .                         | 139        |
| 5.3.5    | Paper-based Presentations . . . . .                                   | 144        |
| 5.3.6    | Handwriting Recognition . . . . .                                     | 144        |
| 5.4      | Linking Documents . . . . .   | 150        |
| 5.4.1    | Model of Links . . . . .  | 151        |
| 5.4.2    | Association Gestures . . . . .  | 152        |
| 5.4.3    | Association Areas for Defining Scopes . . . . .                       | 155        |
| 5.4.4    | Following Hyperlinks . . . . .  | 158        |
| 5.4.5    | Collaboration . . . . .   | 159        |
| 5.5      | Tagging Documents . . . . .   | 161        |
| 5.5.1    | Tangible Tagging with Stickers: Digital Paper Bookmarks . . . . .     | 162        |
| 5.5.2    | Tagging by Association: Tag Menu Card . . . . .                       | 170        |
| 5.5.3    | Tagging with Buttons . . . . .  | 172        |
| 5.5.4    | Tangible Tagging of Processes . . . . .                               | 172        |
| 5.5.5    | Comparison . . . . .  | 176        |
| 5.6      | Conclusions . . . . .   | 176        |
| <b>6</b> | <b>Implementation</b> . . . . .                                       | <b>181</b> |
| 6.1      | Document Representations . . . . .                                    | 181        |
| 6.1.1    | Printed Documents . . . . .   | 181        |
| 6.1.2    | Digital Document Viewers . . . . .                                    | 183        |
| 6.2      | Pen-enabled tabletop display. . . . .                                 | 184        |
| 6.3      | Physical Interaction Tracking . . . . .                               | 186        |
| 6.4      | Pen Data Processing . . . . .   | 186        |
| 6.5      | Annotation Clustering . . . . .                                       | 187        |
| 6.5.1    | Online Algorithm . . . . .  | 187        |
| 6.5.2    | Evaluation . . . . .  | 188        |
| 6.6      | Data Persistence . . . . .  | 190        |
| <b>7</b> | <b>User Studies</b> . . . . .   | <b>191</b> |
| 7.1      | General Design . . . . .  | 191        |
| 7.2      | Lecture Notetaking . . . . .  | 193        |
| 7.2.1    | Method . . . . .  | 193        |
| 7.2.2    | Results and Discussion . . . . .                                      | 194        |
| 7.3      | Reviewing Annotations . . . . .                                       | 198        |
| 7.3.1    | Method . . . . .  | 198        |
| 7.3.2    | Results and Discussion . . . . .                                      | 199        |

---

|          |  |            |
|----------|--|------------|
| 7.4      | Integrating Information from Printed and Digital Documents . . . . . | 203        |
| 7.4.1    | Method . . . . .   | 203        |
| 7.4.2    | Results and Discussion . . . . .                                     | 205        |
| 7.5      | Conclusions . . . . .  | 208        |
| <b>8</b> | <b>Conclusions</b>   | <b>211</b> |
| 8.1      | Summary . . . . .  | 211        |
| 8.2      | Directions of Future Research . . . . .                              | 215        |
|          | <b>List of Figures</b>   | <b>219</b> |
|          | <b>List of Tables</b>  | <b>223</b> |
|          | <b>Bibliography</b>  | <b>225</b> |
|          | <b>Index</b>   | <b>243</b> |

# 1

## Introduction

### 1.1 Motivation

*The paperless office is a myth (...) because (people) know (...) that their goals cannot be achieved without paper. This held true over thirty years ago when the idea of the paperless office first gained some prominence, and it holds true today at the start of the twenty-first century. We hope to have shown that it will hold true for many years to come.*

A. Sellen and R. Harper, 2003 [SH03]

With these words, Sellen and Harper conclude their seminal analysis on how paper is used in today's knowledge work. During the last decades, many attempts aimed at replacing paper in knowledge work. Desktop computing, word processing, electronic mail and the World Wide Web have been considered to have a large potential for replacing paper, which has been a symbol of old-fashioned technology. However as pointed out by Sellen and Harper, the numerous predictions of the paperless office, where paper is completely replaced by digital technology, have not become reality. Despite the advances in computing, traditional paper is still omnipresent in offices, in the field of learning and at home. Paper consumption is even augmenting [SH03, p. 10 sqq.]. Knowledge workers typically work both with paper and digital documents, often simultaneously.

#### 1.1.1 The Gap between Paper and Digital Information

The fact that paper keeps being used cannot be merely attributed to purely technical issues like screen size, resolution and contrast, which were subject to significant improvements in the past and which will continue to evolve. Research shows that, particularly for reading, paper has inherent advantages over digital technologies.

To state only some of these advantages, annotating paper documents with a pen is intuitive, flexible and smoothly integrated with reading. An individual sheet of paper is light, thin and malleable. Therefore, the reader can locate the sheet in a position that is comfortable for reading instead of adapting one's own body to the fix position of a screen. Moreover, paper physically embodies information, making it tangible. The user can utilize both hands to interact with the information. Amongst others, this provides for efficient navigation within documents, for example when thumbing through a book and sensing the appropriate number of remaining pages with one finger. In contrast, navigation in digital documents typically relies on visual information only. Paper also can be flexibly arranged in the physical space. This supports for example sorting and structuring tasks that are carried out by displacing physical sheets of paper. Further, activities that involve cross-document use, such as comparing pages or reading a document and taking notes on another one, can be efficiently carried out with paper. These are only some examples of the advantages of paper, which we will identify in more detail in the next chapter. Empirical work further indicates that working with paper does not only imply a lower cognitive load than working with traditional computers, but it also has lower cognitive load than working with a digital stylus on a Tablet PC or a graphical tablet [OAC06].

Despite these advantages of paper, it is a matter of course that digital media have other, equally important benefits. For instance, digital documents can be efficiently searched, archived and shared over distance. Moreover, they can include dynamic and interactive contents, such as audio, video or data that reacts on user input or other influences.

Due to these unique benefits of both worlds, knowledge workers typically use both paper and digital media. Depending on the type of information and the context, some information is preferred in a printed form while other information is accessed using digital technology. For example, paper might be preferred for reading a longer document, while a computer is used for composing new documents or for gathering information on the Web.

This combined use of printed and digital documents leads to disruptive transitions. The user has to get by with different representational media as well as with different interactions and tools. Most important, while many digital documents can be easily printed on paper, there is no efficient way back to digitize information which has been added on paper. It is therefore the goal of this thesis to reduce the gap between printed and digital documents. A tight integration of both media presents novel opportunities for improving computer support for knowledge work.

### 1.1.2 Knowledge Work with Documents

Working with documents is a – possibly even *the* – central activity in knowledge work. In this thesis, we are particularly interested in knowledge work with existing documents as opposed to a focus on composing new documents. Existing documents are an essential means for acquiring new knowledge, a process that is

typically done by reading and understanding documents and by integrating their contents with prior knowledge. When dealing with existing documents, knowledge workers do not only read them in a sense of a rather passive consumption. Instead, during active knowledge acquisition, the documents are a working surface for cognitive processes. They do not remain untouched but are subject to modifications and extensions: Knowledge workers annotate documents, integrate contents from various sources by creating references to other passages and tag passages for abstraction and later retrieval. The results from the literature and of field studies which were conducted as part of this thesis show that these extensions and modifications are important parts of successful knowledge work.

Making *notes and annotations* is a substantial part of reading processes. Opposed to what we intuitively understand by reading – deciphering words and phrases and ultimately meaning – reading documents frequently comprises writing as well. Commenting, underlining and highlighting a document during reading supports better understanding, critical thinking as well as remembering the thoughts the reader had. Adler et al. call this process *active reading* [AD72]. As Adler notes, “the physical act of writing, with your own hand, brings words and sentences more sharply before your mind and preserves them better in your memory”. Annotations and notes are not only central to reading but also important for efficiently attending lectures or meetings. Taking personal notes or annotating handouts stimulates attendees to actively follow the lecture, to consciously select important content, and to rephrase it in their own words. Moreover, notes and annotations have a reminding effect during review after class [Kie89].

In addition to active reading, it is crucial for successful knowledge acquisition to structure the entire problem domain, to abstract and to establish relationships between concepts. The resulting structural knowledge facilitates recall and comprehension and is essential to problem solving [JBY93]. For this purpose, *linking and tagging* documents are highly relevant activities. These can transform an unsorted and possibly confusing collection of a large number of disparate documents to a unified and well-structured document space.

Annotating, linking and tagging are rather *generic activities* that apply to a broad variety of document-based knowledge work. While there are well-established practices for annotating, referencing and structuring paper documents, these activities are harder to perform with current technology for digital documents. In particular, the transitions between printed and digital documents are not well supported. Notes and annotations made on paper cannot be easily digitized and it is hard to create references between printed and digital documents. Moreover, while it has become common to tag documents on the Web, these systems cannot be used for tagging content that is available on paper.

A further aspect which is essential for gaining new knowledge is *collaboration*. By the exchange with other people, knowledge workers can gain new insights and perspectives, can critically examine their own understanding and can co-construct a shared understanding with others. Even though traditional paper is well suited for many types of co-located collaboration, it obstructs remote collaboration, as it is

more difficult to share physical contents over distance than digital information.

## 1.2 Contributions of this Thesis

This thesis is situated at the intersection of human-computer interaction, computer-supported collaborative work and technology-enhanced learning. It addresses the problems and limitations of current practices of knowledge work which have been discussed above. The overarching goal of this thesis is to provide better technological support for collaborative knowledge work which is based on both printed and digital documents.

During the past few years, a new area of research has formed that deals on the integration of printed and digital information. Real paper is digitally augmented in order to become an interactive digital surface. This area can be assigned to the fields of Tangible User Interfaces and Ubiquitous Computing, which aim at extending computing beyond the desktop into the physical space that surrounds us. Physical objects of the everyday life become key parts of digital user interfaces and act as instruments for interacting with computers. Previous research on the integration of paper and digital information has mainly focused on technology and on individual interaction techniques but has rarely taken on an integrated viewpoint on the various activities that are central for successful knowledge work. Moreover, most of the projects realized so far bridge the gap between printed and digital documents by transferring *information* from paper to computers or vice versa. However, the *interaction* is not tightly integrated, as almost all approaches require the user applying different interaction metaphors and switching between different interaction devices for printed and digital documents.

In this thesis, we show that by following an observation-driven and empirically-inductive research method, we can take the interaction with printed and digital documents one step further. The scientific contributions of this thesis are in the fields of interaction theory, interaction techniques and interactive systems. They are summarized in Fig. 1.1.

**Field studies.** The contributions of this thesis are grounded on *empirical research*. In three initial field studies, we have explored how paper and digital media are used in learning at universities. These field studies were both quantitative studies with several hundred participants and contextual inquiries with students in their real learning environments. The results of these studies allow a *better theoretical understanding of paper-based knowledge work*. They provide quantitative evidence on the use of paper and laptops in lectures and on the influences of these different media. Moreover, they detail on the interwoven character of printed and digital documents in many settings of knowledge work and identify functions and roles of documents as well as typical interactions that are performed with paper. Based on these studies, we identify main activities of knowledge work with documents and show that an *ecological perspective* is an appropriate theory for guiding our work.

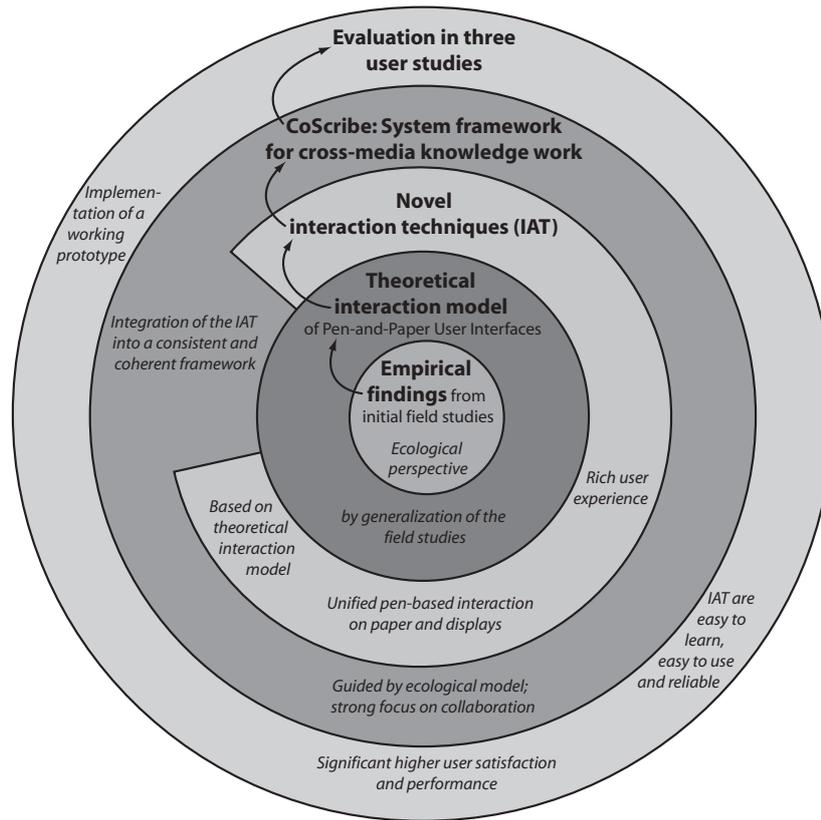


Figure 1.1: Contributions of this thesis

**Theoretical interaction model of PPUIs.** In an inductive empirical process, we further generalized these findings and contribute a *theoretical model of Pen-and-Paper User Interfaces (PPUIs)*. Prior research in this field has focused on systems and not on theory. The model is therefore an important contribution for the theoretical understanding of digital interaction with pen and paper. This model provides general guidelines on how to construct simple and intuitive PPUIs. Its underlying principle is a systematic separation between a semantic and a syntactic level of interaction. Based on this separation, syntactic interaction primitives are identified, which act as building blocks for PPUIs that support specific semantic activities.

**Interaction techniques.** Furthermore, we contribute *novel interaction techniques and visualizations* for cross-media knowledge work with documents. They focus on the cohabitation [Gui03] of paper and computers, i.e. both media are used in combination and treated at an equal footage. These techniques and visualizations enables users to annotate printed and digital documents, to integrate documents with hyperlinks and to abstract from contents to higher-level concepts with tags. In terms

of functionality, this is similar to Bush's [Bus45] vision of Memex, a machine that supports reading and learning processes which are based on documents. In contrast to Memex, the user cannot only work with documents which are displayed by a machine but also with printed documents. These interaction techniques offer a *rich user experience*, being inspired by the traditional practices of using paper and relying on such varied interactions as writing on paper, arranging several sheets of paper, connecting paper sheets and attaching physical stickers. This varied user interface stands in contrast to many previous approaches that leverage only the interaction of writing with a pen on paper. In addition, the interaction with paper and digital information is more seamlessly integrated than in previous work, since the same digital pen and the same interactions can be used both on printed and on digital documents. Nevertheless, the flexibility and mobility of paper is maintained. We call this concept *unified pen-based interaction*. On a technical level, this leverages digital pens that capture and digitize the pen strokes made on paper. In order to support the same interactions on digital documents, we developed a screen prototype that provides for capturing input with these pens.

**System framework.** Moreover, we contribute *CoScribe*, a system framework for collaborative cross-media knowledge work. CoScribe includes the novel interaction techniques and visualizations developed in this thesis. It proposes an *integrated solution* instead of support for partial aspects only as it is done by most prior work on the integration of paper and digital information. Following an ecological model of knowledge work as our guiding theory, CoScribe covers entire workflows in knowledge work with a strong emphasis on *collaboration*. Several persons can work at the same place using multiple pens or can collaborate over a network connection.

**Evaluation.** CoScribe and the novel interaction techniques and visualizations were evaluated in three further *user studies*. For evaluation purposes, we implemented a *working prototype* of the CoScribe system. The results of the evaluation provide evidence that CoScribe significantly enhances both work performance and user satisfaction. They demonstrate a significant performance gain for an information integration task in hybrid collections of printed and digital documents. The studies further show that the new interaction techniques are easy to learn, easy to use and reliable.

### 1.3 Research Method

This thesis was performed in an interwoven process that combined a cyclic design process with an inductive-empirical process for generating theoretical results (Fig. 1.2).

The design process followed the principle of user-centered design [ND86] [Bai04, p. 763 sqq.]. User-centered design is a design philosophy that pays extensive attention to the needs of the end users. Involving end users in the design process

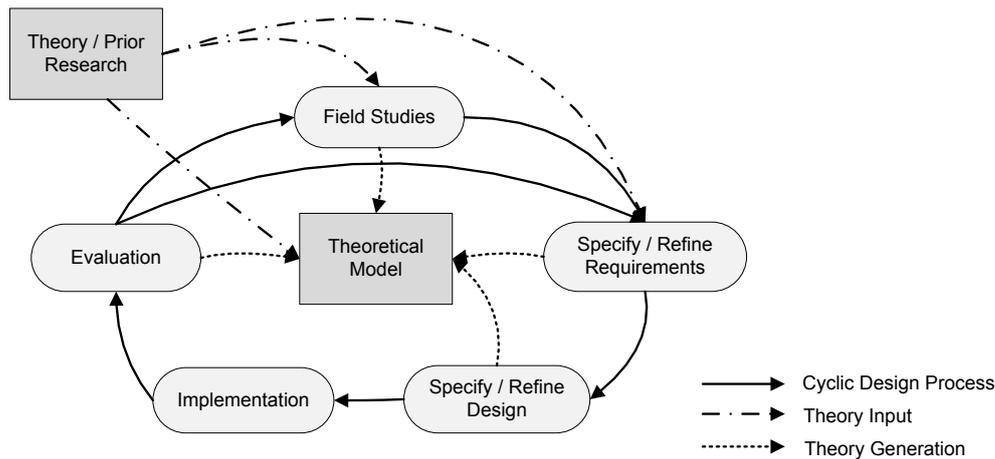


Figure 1.2: Schematic overview of the research method used in this thesis

is crucial for successfully designing a product that corresponds to the needs and wishes of the people it is made for. In many aspects, end users are experts for the domain they are working in. Experience shows that design processes which do not involve this expertise run the risk of imposing a non adequate work style on the users which is defined by external non-experts. Moreover, there are often hidden qualities in work practices, which can only be identified by a thorough examination of the field. For example, at first glance it appears reasonable to share documents electronically within an organization in order to avoid wasting paper with print-outs. Yet, ethnographic research has shown that printing a document and personally handing this physical copy over to a co-worker is preferred by many people to sending an e-mail, as it allows providing additional information or insisting on the urgency in a personal conversation [SH03, NHH06].

For this reason, our design process included extensive empirical work. As depicted in Fig. 1.2, it was performed in an iterative and incremental development cycle [Gra89, ISO99]. Each iteration started with initial field studies. These lead to new or refined requirements. We then designed and implemented novel interaction techniques and visualizations before these were finally evaluated with end users in order to test the validity of the assumptions. Within a two-years period, the project ran through two main cycles and several sub-cycles (see Fig. 1.3). For the sake of readability, this thesis integrates the results of all cycles and presents them in a linear manner.

End users were involved in the field studies and the evaluations, but also contributed to the design process, for instance by testing paper mockups and early prototypes. These end users were university students enrolled in different years of study in various disciplines (including computer science, pedagogy and psychology). Moreover, scientists of our own and of other departments participated to the design process. A varied set of methods was used for gathering information

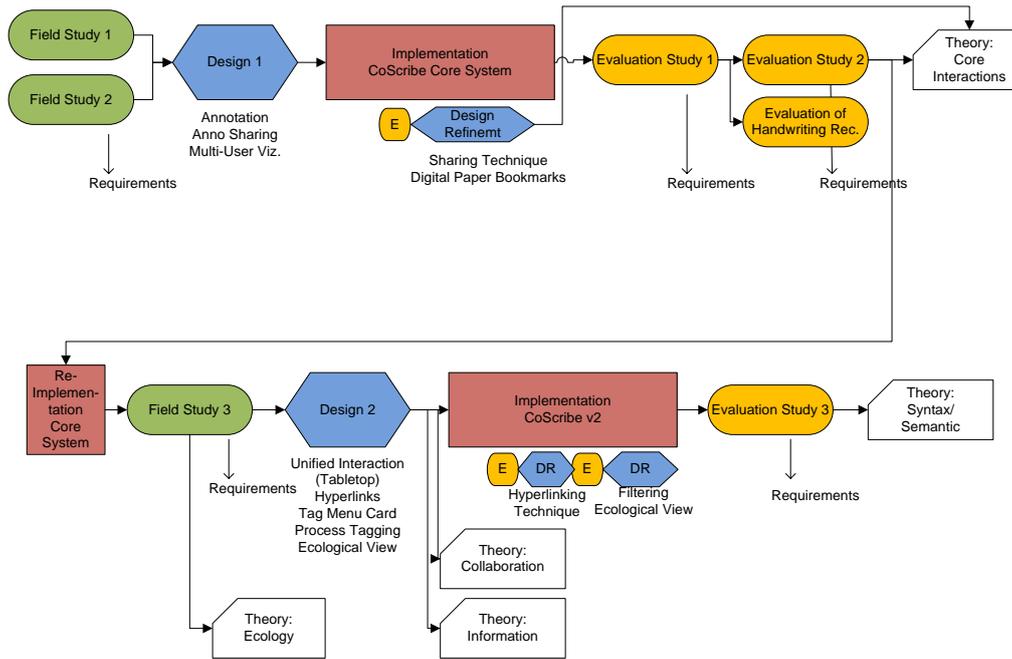


Figure 1.3: Temporal overview on the phases of the research process

from end users. This includes the use of questionnaires, ethnographic observations, semi-structured interviews with individuals and groups of users, discussions of paper mockups, controlled usability testing in our laboratory and analyses of artifacts created by users with our new prototypes.

A parallel theoretical process provided theoretical input to the different phases of the design process. This input consisted of theoretical models of knowledge work and human-computer interaction, of empirical results from the literature and of design solutions from related work. In the reverse direction – from the design process towards theory – we subsequently abstracted the results of the design process in an inductive-empirical manner. This led to the theoretical model of Pen-and-Paper User Interfaces and paper-based document work.

## 1.4 Publications

Parts of this thesis are published in journal articles and in proceedings of international and national conferences and of a workshop. The field studies have been addressed in [SGM07, SB08]. The interaction model of Pen-and-Paper User Interfaces is published in [Ste09]. The interaction techniques for collaborative annotation have been presented in [SBM09a, SBM08a]. Cross-media hyperlinking is addressed in [SBM09b, SBM08b]. Finally, techniques for bookmarking and tagging have been published in [SBM08c, SBM08d].

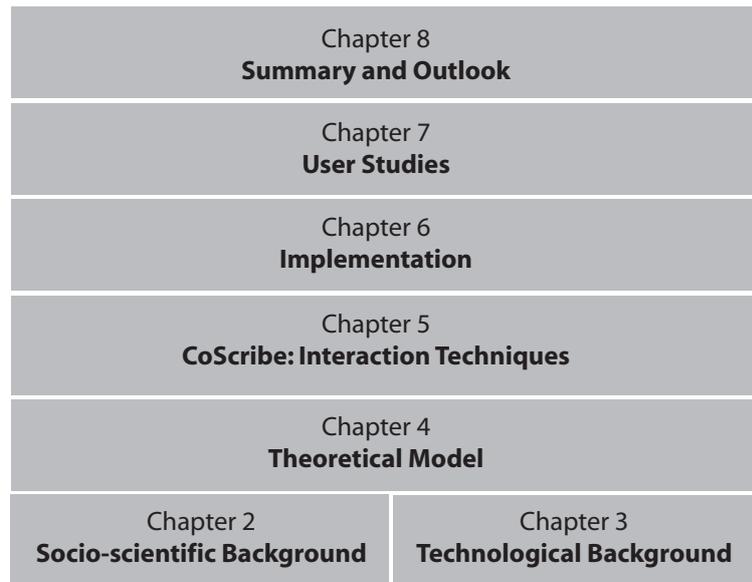


Figure 1.4: Structure of this thesis

## 1.5 Structure of this Thesis

Figure 1.4 gives an overview of the structure of this thesis, which is as follows.

Chapter 2 and Chapter 3 provide a basis for the remainder of this thesis. In *Chapter 2*, we approach the field of paper-based knowledge work from a socio-scientific perspective. In order to familiarize the reader with the specifics of paper, we review and systematize results from the literature on the affordances of paper documents. Further, we complete these general findings by the results of three field studies which were conducted as part of this thesis. These investigate more closely the use of paper and digital media in our field of application, which is learning at universities. Their results bring us to choose an ecological perspective on interaction as our guiding theory, which is presented in the following. Finally, based on the findings discussed in this chapter, we identify a set of requirements for systems which support cross-media knowledge work with documents.

*Chapter 3* discusses the state of the art of paper-based user interfaces. We provide a systematic overview and give a critical analysis of existing work. On the one hand, this comprises theoretical models of paper-based user interfaces. On the other hand, we discuss core technologies that enable to digitally augment real paper as well as interaction concepts and systems that provide for annotating, linking and tagging documents on paper.

In *Chapter 4*, we present a theoretical interaction model of Pen-and-Paper User Interfaces, which covers the three dimensions of interaction, information and collaboration. The model is empirically grounded on the results of our field studies. It

models how information is distributed and related across paper and digital media and how paper supports collaboration amongst multiple users. The main focus of the model is on the interaction with paper-based user interfaces. We identify a set of generic interaction primitives that can be performed with digital pen and paper. It will be demonstrated that systems from related work can be classified in terms of these interaction primitives. However, the model is not only analytical but also generative and provides guidance for the design of Pen-and-Paper User Interfaces. By the combination of several interaction primitives, complex paper-based user interfaces can be designed that offer a rich user experience while nevertheless being easy to use and reliable. This theoretical approach is at the foundation of the novel interaction techniques, which are presented in the following chapter.

*Chapter 5* presents the CoScribe concept for cross-media knowledge work with documents. After an overview on the entire approach and its main components, three sections detail on the concrete interaction techniques and visualizations. These aim at supporting the collaborative annotation, linking and tagging of printed and digital documents. We contribute a flexibly adaptable paper-based user interface for annotating documents and novel means for asynchronous collaboration around paper documents. This includes a paper-based mechanism for classifying and sharing annotations, which is seamlessly integrated with annotating and notetaking. For the review of shared annotations of other users, we propose a novel visualization that integrates both one's own and shared annotations in one single view. This is particularly supportive for gaining an overview on all annotations and for seeking specific annotations. We further discuss how paper-based annotations can be used in lectures not only by the attendees but also by the presenter. Finally, we evaluate the performance of handwriting recognition with free-form annotations made on lecture slides. The rather low performance indicates that in this domain, paper-based interaction techniques should be designed not to exclusively rely on handwriting recognition.

We further contribute an interaction technique for creating and following hyperlinks between printed and/or digital documents, such as Web pages. The underlying interaction metaphor of this technique is a pen-based association which crosses the boundaries of individual display surfaces and of paper and screens. Hyperlinks can be shared with co-workers and are automatically visualized not only within the documents but also in a collaborative view of all users and all documents. Finally, we propose four novel techniques for tagging documents and temporal processes. These are based on different interaction metaphors: attaching tangible index stickers, using association gestures on physical tag cards, tapping on printed buttons and manipulating the orientation of shared, tangible tools. This underscores the variety of interaction styles and the richness that is possible with physical paper.

The CoScribe concept and the interaction techniques were implemented in a working prototype. In *Chapter 6*, we provide an overview on the software and hardware components of CoScribe and describe some important aspects of the implementation that can be generalized beyond this working prototype.

*Chapter 7* presents the results of three user studies of CoScribe. These assess the

use of CoScribe for taking notes and making annotations during lectures, for reviewing own and shared annotations and for integrating information from hybrid collections of printed and digital documents.

Finally, in *Chapter 8*, we summarize the outcomes of this thesis and provide a critical analysis of the results. We also discuss directions of future research in the field of cross-media interaction with paper and digital information.



# 2

## Knowledge Work with Paper Documents

In this chapter, we will approach the field of knowledge work with documents. We will thereby adopt a socio-scientific point of view and explore how people interact with documents – both with traditional paper documents and with digital documents. Based on these findings, we will derive implications for the design of digital systems which support knowledge work with documents.

This chapter is structured as follows. In Section 2.1, we define knowledge work and concretize our application scenario, which is learning at universities. We show that this includes highly varied working methods and focuses on the reception and integration of documents. As the paper medium takes an important place in our application scenario, we then investigate the specifics of paper. In Section 2.2, we review of the affordances of paper from the socio-scientific literature. This review shows that paper offers specific advantages that current digital technology cannot provide in this form. This includes easy navigation, intuitive annotation, flexible organization of information in the physical space, mobile use and strong support of collaboration and mutual awareness. In Section 2.3, we present results from three field studies that were conducted as part of this thesis with the goal to examine the use of paper more specifically in learning at universities. Amongst others, these results show that paper is largely preferred for taking notes in lectures compared to digital notes. The use of laptops constrains notetaking, leading to a significantly lower number of annotations compared to pen and paper. Further, an ethnographic field study of learning group meetings enables us to identify three main functions of documents and shows that learners frequently work with hybrid collections of printed and digital documents, which form an interconnected information space. In addition, we point out the importance of spatial arrangements for organizing information and show that the interaction between learners is most often grounded on interactions with documents, such as pointing to a document or

handing a document over to a co-worker. The insights of these studies lead us to take on an ecological perspective of knowledge work as the guideline for developing the theoretical basis of this thesis. This perspective is described in Section 2.4. It takes on an integral viewpoint on the ensemble of collaborating users, of physical and digital artifacts, of work practices and of their interplay. Amongst others, this perspective resulted in a strong emphasis on collaboration support, on support of physical activities and on support for relating and integrating multiple documents. In Section 2.5, we derive requirements for the design of digital systems that support document-based knowledge work. We postulate that systems should provide support for both printed and digital documents, should enable users in annotating, linking and tagging documents and should support various forms of collaboration. Moreover, novel practices should be easy to integrate into current ecologies. Finally, we draw conclusions in Section 2.6.

## 2.1 Scenario: Learning and Knowledge Work at Universities

This thesis examines and designs technological support for knowledge work. The concept of the knowledge worker was first coined by P. Drucker [Dru64]. As early as in the mid-20<sup>th</sup> century, he argues that knowledge has become the key resource of our society. As opposed to traditional workers who perform manual labor, a knowledge worker works primarily with information. Drucker predicted that knowledge workers will become the largest working group. Following Drucker, we define *knowledge work* as a category of work which primarily deals with using and developing information. Some very clear examples of knowledge workers are scientists, university students, librarians, lawyers, teachers and engineers. Knowledge workers spend a lot of their working time with documents.

In this thesis, we focus on a specific scenario of knowledge work, which is learning at universities. In this context, we define *learning* as a process of knowledge acquisition. Learning at universities is both our scenario of analysis and the scenario we tailored our design solutions for. Nevertheless, our results can be, to a certain extent, transferred to other knowledge work settings, for example to scientific research or to industrial planning, research and development. This is possible because learning includes many rather generic activities which are also part of other knowledge work settings. This includes for example the activities of reading, annotating, writing, referencing, memorizing, summarizing as well as discussing and sharing with other people

Learning at universities is marked by a very intense work with documents and comprises a large variety of working forms. Some examples are depicted in Fig. 2.1. In a first dimension of *guidance*, learning activities can be located on a scale ranging from guided learning to highly autonomous learning. Attending a lecture and reviewing the contents after class is a rather guided form of learning. Individually researching literature, excerpting documents, integrating and abstracting the information found and finally composing a term paper is far more autonomous. In a

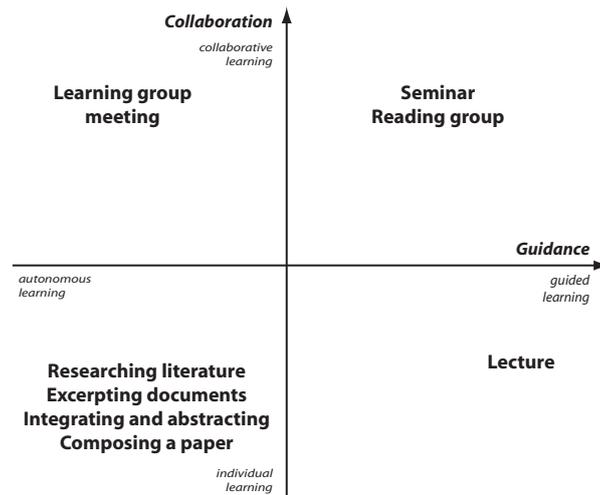


Figure 2.1: Two-dimensional space of learning forms

second dimension of *collaboration*, learning at universities presents a great diversity as well. In addition to individual learning, students collaborate in learning groups meetings, reading groups and seminars. This is mostly done in a co-located manner, but due to the advances of technology and to numerous initiatives for fostering technology-enhanced learning, remote collaboration over distance is augmenting. Collaborating with other learners supports students in gaining additional insights, in critically examining their own understanding and in co-constructing a shared understanding with other learners. As a matter of course, there exist smooth transitions between different types of learning and collaboration, which even augments this diversity.

In this context, central activities for acquiring knowledge from documents are reading, taking notes and making annotations. Moreover, efficient learning integrates new knowledge with existing knowledge, for example by relating information and abstracting it to concepts. This generates structural knowledge of how concepts of a domain are interrelated. While during these activities learners also generate new documents, the focus is on the reception, understanding and integration of existing documents.

## 2.2 Affordances of Paper Documents

One of the most important media in knowledge work is paper. For many years, the paperless office, in which paper is not used any more, has been predicted. Yet, many attempts of companies and organizations to banish paper from their offices have failed. Paper is still a key information medium and its use is even augmenting [SH03]. When designing technology with the goal to support knowledge work, one

should therefore be aware of the characteristics and specifics of paper as well as of the well-established practices of using paper. To understand the specifics of the paper medium, we will systematize results from the literature in this section before we then present our own field studies on the use of paper in learning settings.

The concept of affordances has proven to capture very well the essence of the interaction with paper. The term *affordance* was coined by the psychologist J. J. Gibson [Gib77] and later refined by D. Norman [Nor02]. Both authors state that the physical properties of an object define the possibilities for acting on or with this object.

**Definition 1** (Affordance [Nor02]). *An affordance is a quality of an object which the user is aware of and which allows the user to perform an action.*

For example, the affordance of a light switch is its property to have two physical states, which enables users to switch between an 'on' and an 'off' state. The fact that a paper sheet is thin and flexible affords the action of folding.

The affordances of paper are mainly related to its physical nature. The physical nature of paper provides for a direct and intuitive interaction, which is not mediated by technology. It physically embodies information and renders it tangible. Paper can be moved and arranged in the physical space, one can interact with paper using both hands and paper can be easily written on. Moreover, paper is thin, lightweight and cheap.

We systematize the main reasons why paper supports knowledge work so successfully along the following five key aspects.

### 1. Ease of navigation

Paper supports *easily and flexibly navigating* through documents. Navigation through a paper document is quick and somewhat automatic [SH03]. People leverage the tangibility of paper, which embodies information, and use both hands to search and skim through the document and to interleave navigation with other activities, for instance to flip pages while writing with the other hand. Moreover, the physical thickness of a document provides implicit information about the length of the document and the remaining number of pages.

Due to the static layout of paper documents, information is fix with respect to a physical page. O'Hara et al. [OS97] show that users acquire incidental knowledge of the location of information by reference to its physical place on the page, which helps them in finding this information later on.

### 2. Intuitive annotation

As Sellen and Harper [SH03, p. 82] point out, reading occurs with writing more often than it occurs without. In a diary study, the participants combined reading with writing in more than 75 % of the time and up to 91 % of the time. In this context, writing means both annotating the document itself or writing a separate document.

We define annotations as follows:

**Definition 2** (Annotation). *An annotation is an amendment to an existing document that is conceptually separate yet contextually related. It adds an additional layer of information to it leaving the original document unchanged. An annotation consists of a context and of some content. Its context is the portion of the document the annotation applies to. The content of an annotation is the information added to the original document.*

In contrast to an annotation, a *note* is not added to the contents of an existing document. It has some content, but no context. Informally speaking, a note is a comment on an empty document while an annotation is a comment on an existing document. We will not always clearly distinguish between annotation and notetaking because on a technical level, handwritten notes can be modeled as annotations on an empty document.

An important affordance of paper documents is that they can be *easily annotated with handwritten free-form annotations*. These are very flexible and may include highlightings of portions of the document, textual annotations, formulae, sketches, drawings etc. as well as combinations of these. Pen and paper annotations present a great fluidity in form [Mar98] and can both be very informal or more structured, for example if the user follows a specific annotation or notetaking method (e.g. the Cornell Notetaking Method [Pau89]). Further, handwritten annotations on a printed document are clearly separated from the document and several people can annotate the same copy of a document [SH03]. While this is also true for the broader class of pen-based interfaces, including pen interaction on digital displays such as Tablet PCs, there is some evidence that the use of pen-enabled displays generates greater extraneous cognitive load than the more familiar interactions with real paper [OAC06].

### 3. Flexible spatial organization

In addition to interacting with single sheets of paper or with several sheets bound in a single document, paper affords the combined work with several documents and several sheets of paper that can be flexibly arranged. Adler [AGH<sup>+</sup>98] points out that about half of the reading processes include cross-document use. By arranging several documents on the desk, readers dispose of multiple “display surfaces”.

The physical arrangement of document pages enables to *lay out information in space* and provides *rich ways of expressing the functions and priorities of documents as well as their relationships*. Note that even advanced digital technologies, such as Tablet PCs or E-Book readers, cannot provide these affordances, as users typically dispose only of a single or a very small number of displays. The typical workplace of a knowledge worker contains different functional zones for documents of different priorities [SH03]. For instance, “hot” important documents, which are currently being used, are positioned within the user’s center of attention while “warm” documents are placed at the outer zones of the desk, where they do not interfere with the

current task but nevertheless are quickly at hand. “Cold” documents of little relevance are filed in shelves in the same or even in another room. Physical cues also support to express relations between documents, for example by adjusting them on a pile, by placing them into the same binder or by putting them onto the same shelf. Summing up, spatial arrangements are important for expressing and gaining a sense of the overall structure as well as for referring to other documents. Moreover, the layout of the paper sheets is dynamically changed accounting for the activity which is currently performed. For example, reading and writing require placing the paper sheet at a different angle and a different distance from the reader [OS97].

Paper also affords to *seamlessly interweave multiple activities*. This heavily draws on the intuitive and two-handed interaction as well as on the concurrent use of multiple display surfaces. We already stated that navigating within a paper document is frequently interleaved with other activities like reading or writing. Physical arrangements placing two pages one besides the other facilitate interleaving reading with writing on a separate sheet and comparing or integrating information from several document passages [OS97]. The same is true for placing printed documents that provide information around the computer keyboard when writing a digital document.

#### 4. Mobility

A key affordance of paper is its *mobility*. Since paper is thin, flexible and lightweight and does not require a power supply, pen and paper can be used in a huge variety of situations and physical places. Moreover, the mobility of paper affords interactional flexibility when passing physical copies to other people (see next paragraph).

#### 5. Collaboration and mutual awareness

Paper has specific affordances that support collaboration and awareness of the activities of co-workers. First, using paper documents leaves implicit and explicit traces [Mar97, Mar98], which are helpful for subsequent readers [Wol00]. For example, a textbook in a library, which has been used for some years, contains *implicit traces of use*. Amongst others, nagged and stained pages indicate passages that have been read by many borrowers. Annotations made by previous readers or page markers that have been attached to the book are explicit traces. Marshall [Mar97] states that it is precisely for these traces that many students prefer buying second-hand textbooks instead of new ones.

In an ethnographic study of paper use in air traffic control, MacKay [Mac99] shows that paper supports *mutual awareness* in co-located collaborative settings. At a glance, even from a peripheral viewpoint, it is easy to see if a person works with little or many documents and if she is reading or writing.

Finally, the physical nature of paper documents affords the *personal communication* when handing over a document to a co-worker [Mac99, NHH06].

We have shown that paper has specific affordances, which support important

work practices. Paper efficiently supports *reading and authoring documents*, amongst others because it can be flexibly written on and because it provides for the concurrent use of several documents and for the interleaving of activities. Using paper can facilitate *planning and thinking processes*, mainly due to flexible spatial arrangements. And finally, the use of paper documents proves to be supportive for *collaboration and organizational communication*.

As a matter of course, paper is not always the most appropriate medium for knowledge work. Digital documents provide affordances which are somewhat complementary to those of paper. This includes easy storing and archiving of digital information, dynamic and interactive contents as well as quick full-text search functionalities. In addition, albeit it is relatively cheap to produce paper and to print on it, the cost of dealing with paper documents after printing – delivery, storage and retrieval – can be much higher than the respective cost of digital documents [SH03]. Our research therefore aims at offering support for the combined work with paper documents and digital documents in order to complement the affordances and advantages of both worlds.

## 2.3 Field Studies

The studies cited in the previous section provide fundamental insights into the affordances of paper. Most of the findings are based on workplace observations. In order to examine more specifically the role of paper and digital documents in learning at universities and to provide quantitative data on the use of paper and digital media, we conducted three own field studies. In the following, the results of these studies will be presented, which examine media usage, notetaking and annotation processes as well as collaborative practices in learning at universities.

### 2.3.1 Study I: Notetaking in Lectures

The first study assesses student notetaking in university courses. Notetaking in general and more particularly the use of notetaking and annotation software in learning are not well studied [BP05, BK06]. Therefore, we conducted a quantitative investigation in order to derive implications for the design of digital systems that support users in taking notes and making annotations. Our research is three-fold:

1. We evaluated the reasons for the choice of taking notes with a *pen and paper* or with a *laptop*. We then assessed the effects of this choice on the notes being taken and on further review and completion activities.
2. Our hypothesis was that notetaking heavily depends on multiple *context factors*. Therefore, we developed a context model for notetaking in university lectures (see below) and evaluated the influence of several specific context factors.
3. Finally, we assessed note-based *collaborative activities*.

| Discipline       | Respondents | Female | Male   |
|------------------|-------------|--------|--------|
| Computer Science | 316         | 13.8 % | 86.2 % |
| Pedagogy         | 92          | 78.4 % | 24.6 % |
| Overall          | 408         | 28.8 % | 71.2 % |

Table 2.1: Participants of the study

## Method

This quantitative study was based on a questionnaire. Overall, 408 respondents attending lectures in computer science or pedagogy participated to the investigation. Table 2.1 depicts the number of respondents questioned as well as their gender. The participation was voluntary and no compensation was given.

The questionnaire was handed out at the end of a university semester in five courses. We chose four computer science courses, which covered several domains and in which students of different years of studies were enrolled. These courses were a first-year introductory course to computer science, a second-year algorithm theory course and two different courses on computer networks, which are typically attended in the third or fourth year of studies. In order to provide for an interdisciplinary comparison, one course in pedagogy was chosen, in which students of different semesters were enrolled (in average, they were in their fourth semester with a standard deviation of 2.8 semesters).

All courses contained eLearning elements. The computer science courses offered a web-based forum for discussions among the students. Two of these courses were recorded and a video including the slides was offered for download after each class. In all evaluated courses, the instructors used PowerPoint slides, which were made available as downloads before the courses. In the pedagogical course, the instructor additionally provided a pure textual script covering more detailed contents than the slides.

The questionnaire contained 22 closed and open questions related to four topics:

1. *Notetaking behavior*: media used for notetaking, (dis)advantages of those media, types of the notes and the language they are written in
2. *Collaboration and team work*
3. *Course-related information* such as the amount of time invested for the course, the personal interest in the topics and the perceived degree of difficulty
4. *Personal information* about the respondent like sex, field of study, semester and hardware equipment

In the statistical analysis, we investigated correlations between items which were five-point scaled and performed  $\chi^2$ -tests and t-tests to identify significant group differences. All these tests were based on a level of significance of 95 %.

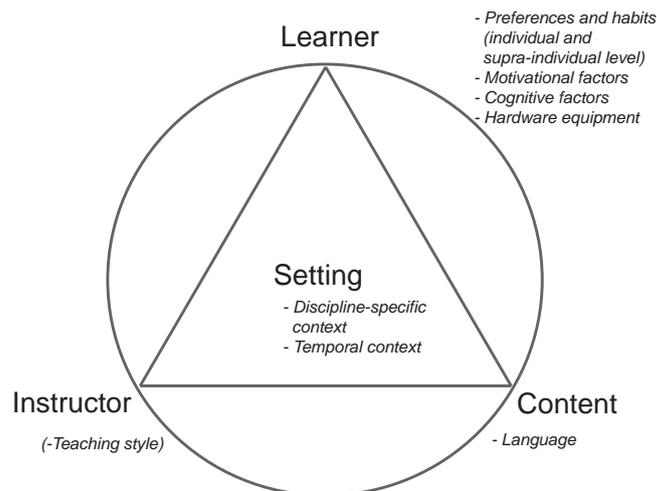


Figure 2.2: Context factors in notetaking assessed in the second field study

### Context Model

In order to systematize our analysis, we developed a model of context factors which are likely to influence notetaking in university lectures. A graphical overview is presented in Figure 2.2. The model puts emphasis on the communicative situation in which notetaking in lectures takes place. It is based on Bühler's *organon* model of linguistic signs and of communication [Büh78]. This model states that linguistic signs (and communication) have three dimensions that are interdependent: the sender, the receiver and the referent the sign represents. Following these three dimensions, we distinguish three central context types (learner - instructor - content). These are denoted by the three corners of the triangle. We add a further dimension which is not covered by Bühler's model. This is the setting that surrounds the communicative situation. It is illustrated by a circle.

We evaluated the influence of several specific aspects of these four context types. The *learner* context type includes preferences and habits, which are personal (and hence on an individual level) or relate to the membership in a social group like gender (supra-individual level). In addition, the influence of two motivational and cognitive factors was assessed, namely the interest in the content and the average grades obtained during previous studies. Moreover, the potential relationship between the ownership of a laptop and the preference for electronic notes was evaluated.

In the *content* type, we analyzed the influence of the course language.

The aspect of the *instructor's* teaching style was not empirically assessed; however, we found a qualitative indice of an influence.

Finally, in the *setting* context type, the study assessed discipline-specific aspects as well as differences between the temporal phases of course review and exam prepara-

ration. We did not include a context aspect of tools used for notetaking (i.e. pen, keyboard etc.), since these are not part of the context but of the notetaking process itself. Furthermore, the types and aspects presented herein are not exhaustive. In further work, this framework can be extended for additional types and aspects of context.

In the remainder of this section, the results of the study will be presented along different categories. We will first describe the groups of respondents taking notes. Next, we will detail the media used for notetaking and the reasons of this choice, particularly with regard to the difference between notes on paper or on a laptop. We will then turn our attention to the notes themselves and discuss their types and the languages they are written in. Finally, follow-up activities of notetaking and collaborative aspects will be analyzed.

### Respondents Taking Notes

The proportion of students taking notes during the course considerably varied between the disciplines. While 93.3 % of the pedagogy students took notes, only 62.3 % of the computer science students did. When asked for the reasons for not taking notes through an open question, the largest group of answers to an open question considers the course slides offered by the instructor to contain sufficient information ( $N = 16$ ). Eight respondents indicated that taking notes distracted them from listening.

A significant difference related to the gender of the respondents was revealed in computer science, but not in pedagogy. While 30.0 % ( $N = 12$ ) of the female respondents in the computer science courses did not take notes, a significantly larger proportion of 48.4 % ( $N = 121$ ) of male students did not take notes [ $\chi^2(1, N = 290) = 4.702, p = .04$ ].

Within the computer science courses themselves, our data reveal a statistically significant difference. In the first-year introductory course, 58.4 % ( $N = 66$ ) of the respondents indicated not to take notes at all, whereas in the other computer science courses, 36.0 % ( $N = 73$ ) did not take notes [ $\chi^2(1, N = 316) = 14.845, p = .000$ ]. We hypothesize that this difference may be related to the minor experience of first-year students.

When relating these results to the context model, we notice that the decision of taking notes or not seems to depend on the setting and learner contexts, namely on the discipline and on the gender of the respondents. The percentage of students who took notes maximally varied from more than 90 % in the pedagogy course to less than 50 % of male students in computer science.

### Media Used for Notetaking

We further asked the students on which media they take their notes. Moreover, we examined the use of the course material (i.e. PowerPoint slides, handouts etc.) and searched for differences which relate to the use of a pen and paper or a laptop.

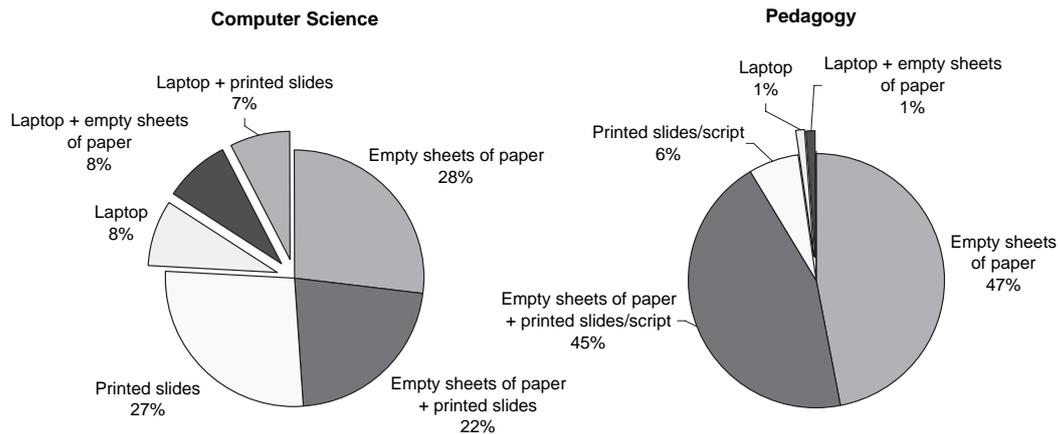


Figure 2.3: Combinations of media used to take notes on

**Pen and paper vs. Laptops.** Figure 2.3 shows the percentage of notetakers on single media or on combinations of several media. Both in computer science and pedagogy, traditional notetaking with a pen and paper clearly outperforms notes on a laptop. In the computer science courses, 77 % of the respondents took their notes exclusively on paper. This group consists of three subgroups of roughly equal percentages which took notes either on empty sheets of paper, on printed course slides or on both of them. 8 % made an exclusive use of a laptop, while 15 % indicated to prefer cross-media notetaking, which combines notes on a laptop with notes on empty sheets of paper or on printed course slides.

The context factor of the discipline proved to be an influential factor of the context model, since laptop use differed largely between the disciplines. In the pedagogy course, laptop use was almost not existent. 98 % took their notes exclusively on paper. The two largest groups (about 45 % each) took notes either only on empty sheets of paper or combined them with printed slides or the printed course script. These findings confirm results of other studies on the choice between paper and laptops [OS97, Obe03], which, however, did not assess notetaking during courses but during overall reading processes and moreover constrained the participants to use a specific software for notetaking.

**Hardware Equipment.** The hardware equipment does not seem to be relevant for the choice of pen and paper. The percentage of students taking notes on a laptop was small even though 78.6 % ( $N = 180$ ) of the notetakers possess a laptop. Only 19.6 % ( $N = 35$ ) of the students owning a laptop took notes on this device.

**Course Specifics.** Comparing different computer science courses, we found that in one course (algorithm theory), the respondents made a significantly higher use of empty sheets of paper (76.4 %,  $N = 42$ ) than the remaining respondents (52.3 %,  $N = 138$ ).

$N = 58$ ) [ $\chi^2(1, N = 166) = 8.927, p = .004$ ]. In this course, the instructor frequently drew sketches and diagrams on the blackboard which were not contained in the slides. We assume that this specific teaching style combined with little free space on the slides led to the heavy use of paper sheets.

**Laptop Notetakers.** Taking is a closer look on laptop users, we asked them about the software they took their notes with. Responses fell under two categories of almost equal frequency: Software that allows to annotate the electronic course slides (e.g. Adobe Acrobat) ( $N = 16$ ) or word processors and text editors ( $N = 17$ ). Four students indicated to use both annotation and a word processor, three students annotated on a tablet PC and two students employed a specific software for creating mindmaps. These data show that the repartition between annotating print-outs and taking notes on blank sheets which we identified for paper notetakers is approximately reflected by notetaking on laptops. The main device for the input of electronic notes is the keyboard, since only few students own a Tablet PC (1.2 %).

**Use of Course Material.** The prominent position which course material provided by the instructor holds in student notetaking is reflected in the general use of course material as well. Figure 2.4 depicts the mean frequency of course material use independently of notetaking. This chart indicates that the most frequently used media are course slides and the textual script.

When comparing computer science students taking notes on paper or on a laptop, our data reveal significant differences in the use of course material. The laptop group made only rare use of printed course slides and of textbooks<sup>1</sup>. Instead, laptop notetakers used the electronic version of the slides very frequently<sup>2</sup>. These data indicate that students tend to avoid media discontinuity between the medium for notetaking and other frequently used course media. However, students making a heavy use of the electronic video recordings did not indicate more often to take notes on a laptop.

### Advantages of Different Media

Besides assessing the distribution among different notetaking media, we aimed to gain information about the reasons for choosing those media as well as the advantages respondents associated with paper or electronic notes on a laptop. Students were requested to judge the importance of several advantages of paper and electronic notes on a five-grade scale. In addition, we posed an open question, in which

<sup>1</sup>Printed course slides: Laptop notetakers  $M_1 = 2.1$  [ $SD = 1.3, N = 15$ ] vs. Paper notetakers  $M_2 = 3.9$  [ $SD = 1.3, N = 147$ ],  $T = 5.133, df = 160, p = .000$ .

Textbooks: Laptop notetakers  $M_1 = 1.5$  [ $SD = .64, N = 15$ ] vs. Paper notetakers  $M_2 = 2.3$  [ $SD = 1.2, N = 143$ ],  $T = -3.77, df = 25.429, p = .001$ .

<sup>2</sup> Electronic slides: Laptop notetakers  $M_1 = 4.9$  [ $SD = .26, N = 15$ ] vs. Paper notetakers  $M_2 = 4.3$  [ $SD = .89, N = 150$ ],  $T = 6.283, df = 59.289, p = .000$ .

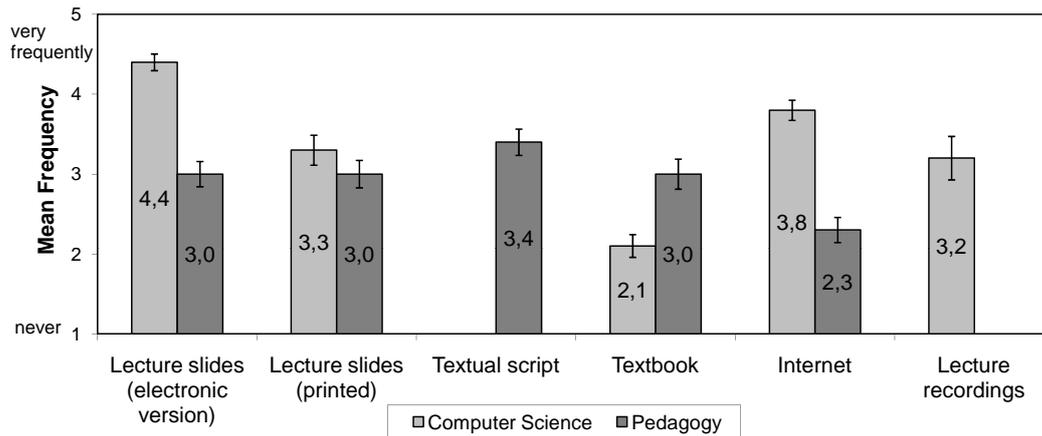


Figure 2.4: Use of course material (independently of notetaking). Error lines indicate the 95 % confidence intervals of the means.

we asked the students to explain why they preferred the specific media they took notes on.

The results of the quantitative question are depicted in Figure 2.5. They show that the free-form flexibility was regarded as the most important advantage of notes on paper. This is followed by the fact that paper can be easily transported. As far as electronic notes are concerned, all proposed advantages were rated almost equally. When investigating differences between the advantages of paper and laptops, we found that long-term archiving was rated significantly more important for laptop than for paper notes [ $T = -5.935$ ,  $df = 234$ ,  $p = .000$ ]. Similarly, good readability of typescript is rated significantly more important than good readability on paper [ $T = -5.907$ ,  $df = 230$ ,  $p = .000$ ].

Students taking notes on paper regarded most advantages of paper as significantly more important than students taking notes on a laptop and vice versa. However, both groups highly rated the flexibility of free-form notes on paper, which thus seems to be of great importance even for laptop notetakers.

The responses to the open question indicate some additional important factors. Students taking notes on a laptop valued that notes can be more easily modified ( $N = 4$ ) and offer a cleaner appearance ( $N = 2$ ). Two respondents stated to prefer electronic notes because this way, they do not have to print the slides. Two further students noted that a laptop allows them to keep the information in one place. On the other hand, 20 respondents stated that notetaking on paper is easier and faster than notetaking on a laptop.

**Annotation vs. Notes.** The responses also indicated reasons for preferring annotations on course material or notes on blank sheets of paper. Annotations on printed course slides are regarded as advantageous, since these allow to establish a direct

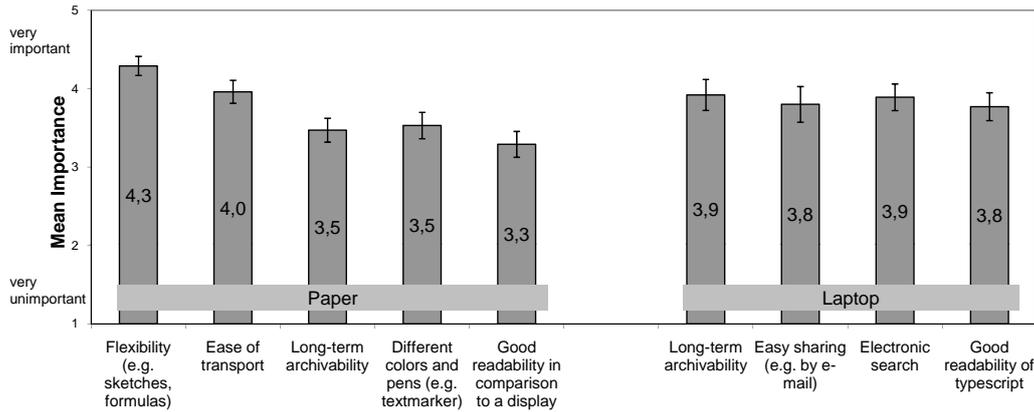


Figure 2.5: Advantages of paper and electronic notes. Error lines indicate the 95 % confidence intervals of the means.

reference to the context by taking the notes on the place they refer to ( $N = 27$ ). 24 respondents particularly valued that they do not need to write everything down on the slide, but only add additional information of importance. On the contrary, blank sheets of paper are favored because they allow to create an own structure and to note own ideas more individually ( $N = 12$ ). Moreover, in contrast to course slides, they provide sufficient free space ( $N = 5$ ). Three students indicated to combine notes on paper and on printed course slides in order to separate their own ideas from additional information given by the instructor.

### Types of Notes

In order to get a broad overview on the notes themselves, we asked respondents what types of contents they noted on paper or with a laptop.

The most frequent notes consisted of writing down or highlighting important points (this was indicated by 81.5 % ( $N = 202$ ) of the respondents who took notes). 39.7 % noted or marked unclear points. 25.4 % took notes containing own questions. A final group of 15.1 % wrote down as many information given during the course as possible. While 50.0 % of the students enrolled in pedagogy reported writing longer sentences, only 23.5 % of computer science students did so.

The choice between laptop or paper does not seem to affect note types, since no significant differences between those two notetaking groups were identified.

As stated above, a selection of important points is made by most of the students taking notes. However, less than half of the students noted unclear points, only one fourth formulated own questions. Hence, a more active note-based work with the material seems to take place rather seldom.

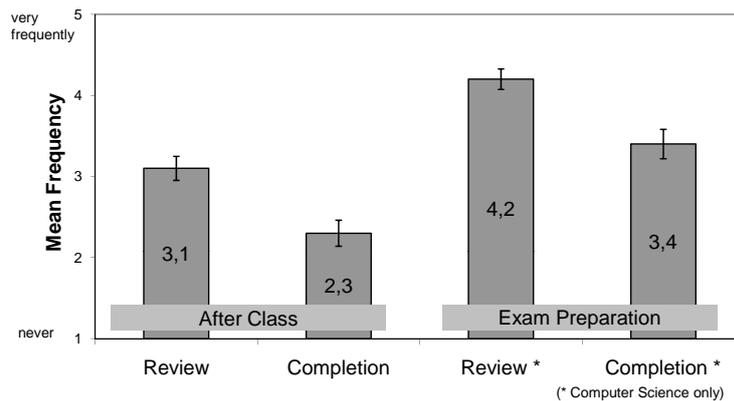


Figure 2.6: Frequencies of follow-up activities of notetaking. Error lines indicate the 95 % confidence intervals of the means.

### Language of the Notes

A further aspect of our context model is the language in which the notes are taken. This aspect must be considered if an eLearning system includes further analysis of textual notes like handwriting recognition or summarization and recommendation of relevant notes.

Even though German stays the most used language, the language in which the course is held largely influences the languages of the notes. The percentage of respondents who indicated to often or very often use the German language varied from 36.7 % ( $N = 14$ ) in courses held in English to 95.0 % ( $N = 207$ ) in courses held in German.

An important finding was that a significant proportion of students combined notes in two or more languages. This percentage corresponded to 60.2 % of the respondents in computer science and to 26.8 % in pedagogy.

### Review and Completion of Notes

Respondents who took notes were asked how frequently they review and complete their notes after class and when preparing the exam. These results allow to estimate in which phases an electronic notetaking system would be used. Mean values are depicted in Figure 2.6. (Results for exam preparation relate only to the computer science courses, since in the pedagogy course, no final exam took place.) The results show that, in contrast to the wrap-up phase after class, where scores are rather low, students become more active when preparing the exam. The context factor of time thus seems to influence notetaking.

No significant differences were found between laptop and paper notetakers. However, taking a closer look on the group which took notes on paper, our data indicate that annotations relate to more frequent follow-up activities than notes on empty

sheets of paper: In review after class as well as in review and completion before the final exam, mean frequencies of respondents annotating printed course slides or the course script were .4 to .6 points higher<sup>3</sup> ( $p < .006$ ) than of students taking notes on blank sheets.

### Note-based Collaboration

In order to assess collaborative behavior in university courses, we further asked the respondents to indicate their collaborative activities which make use of their notes.

54.4 % ( $N = 135$ ) of the respondents who took notes during class indicated to use them for collaborative work. The most important point (71.1 % of this group) consisted of using the notes as a basis for group work and discussion with other students. 51.1 % compared their notes for completion with those of others. 45.9 % gave their notes to other students or used those of others, e.g. in case of illness. Collaborative use of notes does not seem to relate to a specific medium on which respondents took their notes.

Asynchronous notetaking systems offer functions which are comparable to those of web discussion forums. Hence, the use of the forum provided for the course can act as an indicator for collaboration in notetaking systems. 74.7 % of the computer science respondents indicated to use the electronic discussion forum of the course. This score thus outperforms the percentage of collaboration with notes. Within this group, 58.5 % only read the posts. A more active role was taken by 14.5 % who read posts and posed questions and 27.0 % who additionally posted answers to questions of other students.

We found a relationship between collaboration and the frequency of note review and completion. Respondents using their notes collaboratively review them more frequently after class.<sup>4</sup> Alike, this group completes them more frequently after class than non-collaborative notetakers.<sup>5</sup>

### 2.3.2 Study II: Annotating with Pen and Paper vs. Annotating with a Laptop

After having analyzed which media students take notes on as well as the influences of contextual factors, the goal of a second study was to take a closer look on the annotations themselves. In order to assess the influence of the annotation medium on the annotations, we collected annotations made by postgraduate students on presentation slides and manually analyzed this user generated content.

<sup>3</sup>Review after class: Annotators:  $M_1 = 3.3$  ( $SD_1 = 1.1$ ,  $N_1 = 132$ ) vs. Notetakers:  $M_2 = 2.9$  ( $SD_2 = 1.0$ ,  $N_2 = 89$ ),  $T = -2.797$ ,  $df = 219$ ,  $p = .006$ .

Review before the exam: Annotators:  $M_1 = 2.5$  ( $SD_1 = 1.1$ ,  $N_1 = 127$ ) vs. Notetakers:  $M_2 = 1.9$  ( $SD_2 = 1.0$ ,  $N_2 = 88$ ),  $T = -4.034$ ,  $df = 213$ ,  $p = .000$ .

Completion before the exam: Annotators:  $M_1 = 3.6$  ( $SD_1 = 1.2$ ,  $N_1 = 109$ ) vs. Notetakers:  $M_2 = 3.1$  ( $SD_2 = 1.4$ ,  $N_2 = 82$ ),  $T = -2.681$ ,  $df = 189$ ,  $p = .008$ .

<sup>4</sup>Review after class: Collaborative users:  $M_1 = 3.3$  [ $SD = 1.0$ ,  $N = 134$ ] vs. Non-collaborative users:  $M_2 = 2.7$  [ $SD = 1.1$ ,  $N = 92$ ],  $T = -4.142$ ,  $df = 224$ ,  $p = .000$ .

<sup>5</sup>Completion after class: Collaborative users:  $M_1 = 2.4$  [ $SD = 1.0$ ,  $N = 129$ ] vs. Non-collaborative users:  $M_2 = 1.9$  [ $SD = 1.0$ ,  $N = 91$ ],  $T = -3.671$ ,  $df = 218$ ,  $p = .000$ .

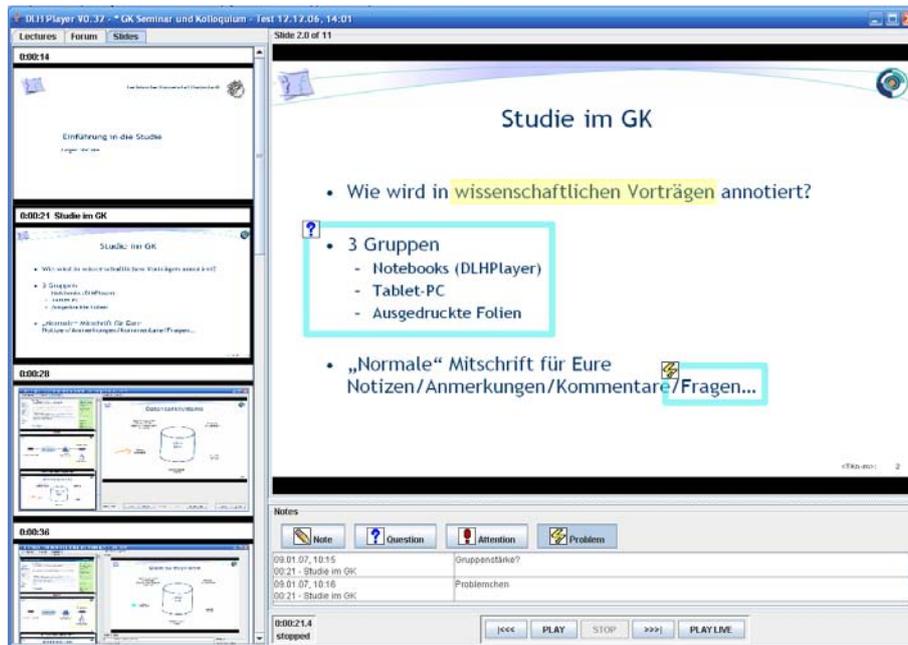


Figure 2.7: Screenshot of the tool for textual annotations and highlightings

### Method

A total of 24 postgraduate students participated to this study. These students attended an interdisciplinary seminar on technology-enhanced learning. Each seminar session consisted of a presentation held by a student and of a subsequent plenary discussion. All presenters used PowerPoint slides during their talks. The participants could annotate these slides. We collected and analyzed annotations from nine seminar sessions, each lasting about 90 minutes and containing an average of 24.7 slides ( $SD = 18.3$ ). We moreover conducted informal interviews after the sessions. Participation was voluntary and no compensation was given. We chose a between-subjects design for the study.

**Pen and paper condition.** One group of students used pen and paper to make personal annotations on printouts of the presentation slides, which we distributed beforehand. In order to guarantee a realistic setting, each printed A4 page contained four presentation slides. (Informal investigations had shown that this layout is used by most students for printing lecture slides.) Figure 2.10 on p. 33 gives examples of handwritten annotations. A handwritten annotation consists of a set of one or more pen strokes a user has made on a page.

**Laptop condition.** A second group of students used an annotation tool developed by our department (see Figure 2.7). This enabled the user to make typewritten anno-

tations and textmarker highlightings on the presentation slides using a laptop. This tool displayed the presentation slides in real-time and moreover enabled manual scrolling through the entire slide set. The tool was designed to make typewritten annotations as easy as possible. In order to make an annotation, the user marks a context area on a slide by drawing a rectangle and then enters the text of the annotation. The context area is then permanently visible. The text of the annotation can be viewed later on by hovering with the mouse on this area. Alternatively, if no text is entered, the context area is displayed as a yellow highlighting in a textmarker style. Individual notes can optionally be labeled with a semantic type (*question*, *attention* or *problem*). Moreover, the user can optionally mark an annotation as public. This annotation is then shared with other users over a network connection and immediately displayed in their slide sets as well.

Students who disposed of an own laptop participated to the laptop group, while the remaining students participated to the pen and paper group. Since not all students attended all sessions, the size of the groups varied. In the pen and paper group, an average of 6 students ( $SD = 1.7$ ) participated to the each session. An average of 4.3 students ( $SD = 2.2$ ) participated to the laptop group. Participants of both groups were instructed to make annotations the same way they normally do during the seminar.

**Annotation segmentation.** A methodological problem with the analysis of handwritten annotations is how to segment all pen strokes on a page into individual annotations. This segmentation cannot always be made unambiguously by persons other than the annotator, since the annotator typically does not explicitly mark the extent of an annotation. For counting the number of handwritten annotations, we had to manually segment them. In order to guarantee a more reliable segmentation than by following merely subjective classification criteria, we used the following rough guidelines for the segmentation procedure:

- An annotation must have a coherent content on a semantic level.
- The strokes forming one annotation must be spatially coherent, i.e. two clusters of strokes that have a large distance are considered as two separate annotations.
- Clusters of strokes that are visually separated in an explicit manner (e.g. by a line drawn by the annotator) are considered as separated annotations.
- Sketches are own annotations. However, if text can be attributed unambiguously to a sketch, both this text and the sketch is considered as one annotation.

In contrast, the segmentation of typewritten annotations was not problematic, as a typewritten annotation consists of a text string. This text can be unambiguously attributed to this annotation.

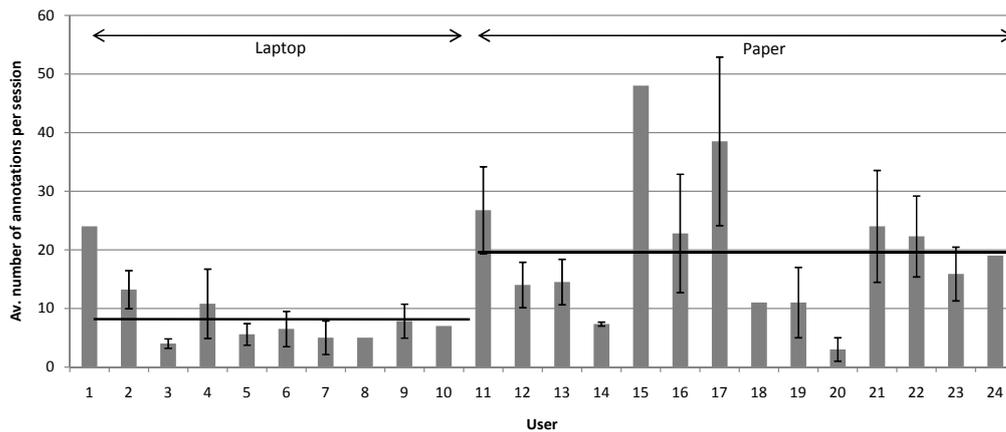


Figure 2.8: Average number of annotations made by each participant per session. The horizontal lines indicate the average over all participants of each condition. Error lines indicate the 95 % confidence intervals of the means.

## Results

We collected a total of 1409 typewritten and handwritten annotations.

A main finding was that the use of a *keyboard seems to constrain the annotation process*. We collected all typewritten and handwritten annotations, manually segmented handwritings into individual annotations which were spatially contiguous and seemed to express the same content and finally counted all annotations. On average, participants of the laptop group made significantly less annotations per participant and session ( $M = 8.9$ ,  $SD = 6.0$ ,  $N = 10$ ) than those of the pen and paper group ( $M = 19.9$ ,  $SD = 12.1$ ,  $N = 14$ ) ( $t = 12.1$ ,  $df = 22$ ,  $p = .015$ ). Informal interviews with members of the laptop group indicate that the participants perceived typing annotations on a keyboard as rather inflexible and complicated. Moreover, making handwritten annotations was judged being faster than typing text on the keyboard. Two participants of the laptop group explicitly stated having observed that they frequently did not make annotations they would make with pen and paper because it took too much effort.

A manual analysis of the annotations showed that the *annotation style is highly individual*, a finding which is consistent with studies of textbook annotations [Mar97, Mar98]. On the one hand, this concerns the number of annotations made. Figure 2.8 depicts the average number of annotations made by each participant per session. In the laptop group for instance, the number of annotations per participant and session varies from 4 to 24. In the paper condition, this number varies from 3 to 48. On the other hand, the style of individual annotations can highly vary. While some users prefer making their annotations within the slides, others obviously prefer notes at the outer zones within the margins. Some participants write short and abbrevi-

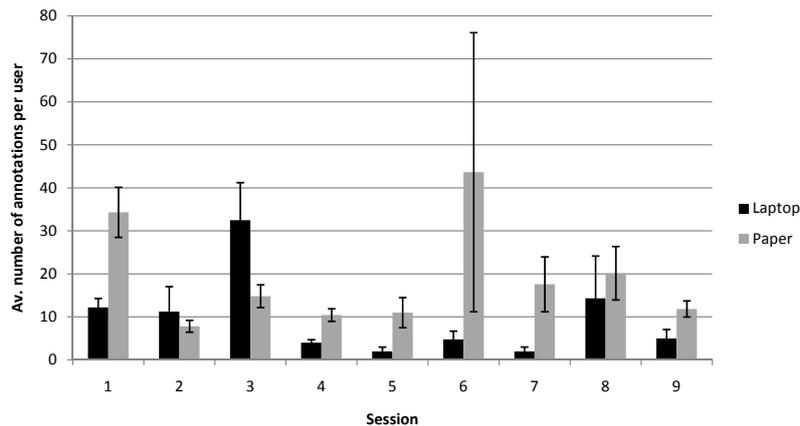


Figure 2.9: Average number of annotations per user of each session. Error lines indicate the 95 % confidence intervals of the means.

ated notes in a telegraphic style or do only underline important passages without making textual comments at all, whereas others tend to write more extensive notes. Figure 2.10 shows examples of different annotation styles.

Another finding was that the *course context* heavily influences the number of annotations made. Figure 2.9 gives the average number of annotations per user for each session. It shows that there is a large variance between the individual sessions.

A further relevant outcome was that the majority of the participants used specific *classification schemes* to classify document contents for later retrieval. In order to do so, participants used symbols with a meaning which is generally well-understandable, such as question marks and exclamation marks (76 % of the participants). Moreover, 43 % of the participants used additional symbols which they have defined themselves and whose meaning would not be evident for others. Figure 2.10 (lower right) shows an example of a user-defined symbol.

We manually analyzed 1097 handwritten annotations and identified types of annotations and their repartitions. Figure 2.11 depicts the type of the annotations. 60.3 % are annotations with an implicit context area, i.e. without an explicit highlighting of the area in the document the annotation applies to. 15.9 % of the annotations contain only highlightings (e.g. an underlining of a phrase or a circle or box around a printed passage). 11.5 % of the annotations contain both a highlighting of the context area and a textual content. Further, 10.8 % are specific annotations that contain only a symbol and act as semantic markers in the sense described above. For this purpose, the use of standard symbols (like question and exclamation marks is more common than using personal user-defined symbols. Finally, 1.5 % of the annotations are lines that are drawn to make a clear delimitation between two or more annotations written near each other.

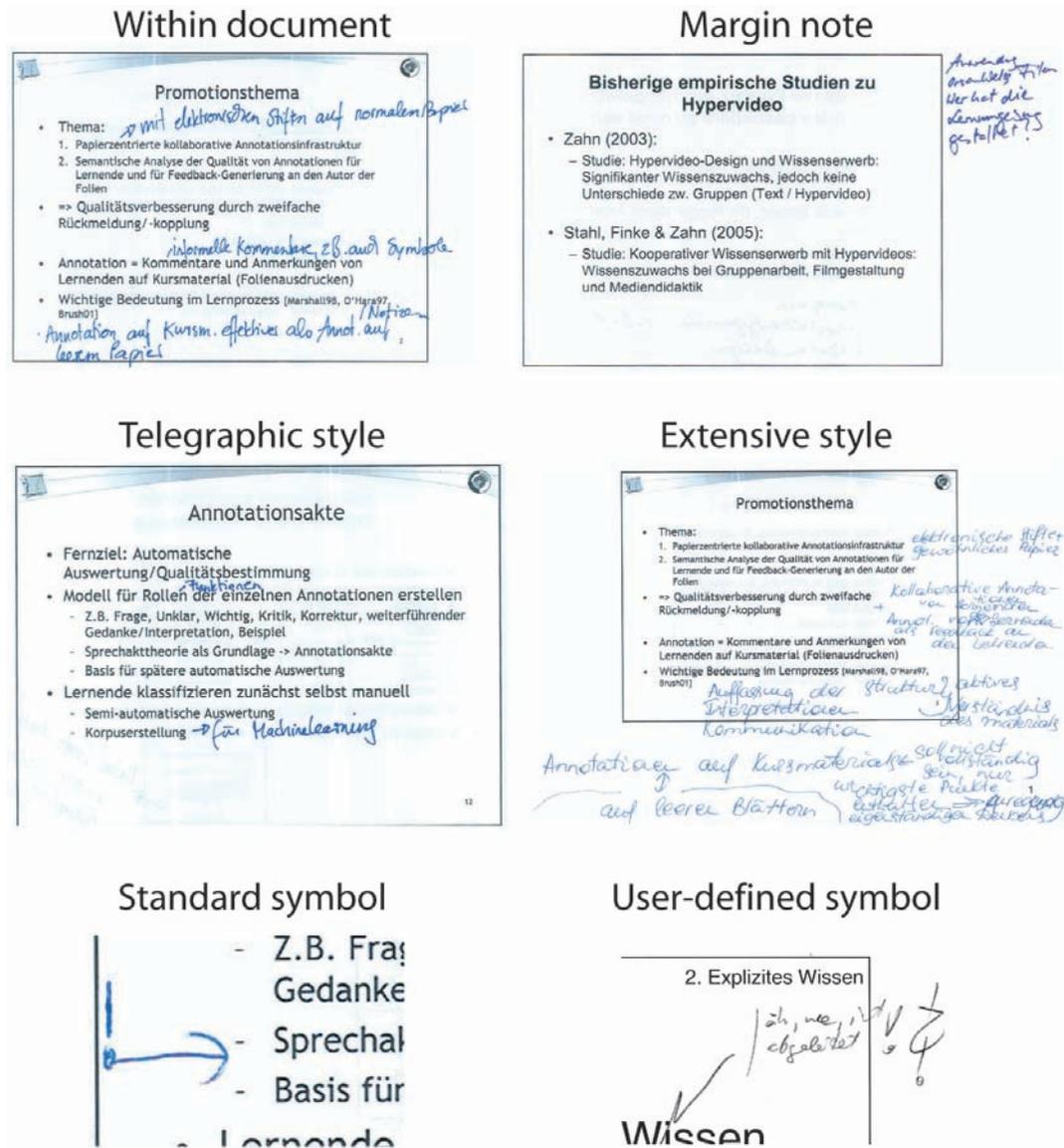


Figure 2.10: Examples of handwritten annotations

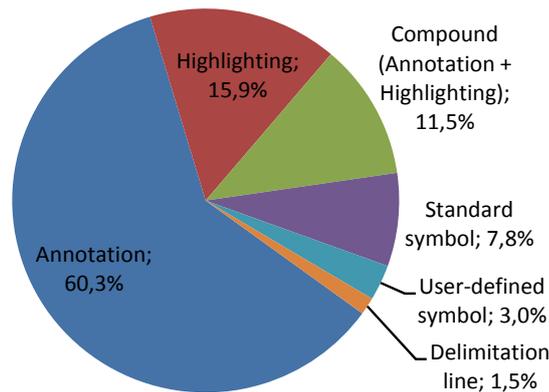


Figure 2.11: Types of handwritten annotations ( $N = 1097$ )

### 2.3.3 Study III: Media Use and Collaboration in Learning Groups

In the two previous studies, we had only analyzed the individual annotation behavior. In an ethnographic study, we therefore focused on the collaborative use of lecture documents in learning group meetings. The goal of this study was to find out which document types are frequently used, to analyze the functional roles of these documents and their relations and to explore the different ways how participants interact with documents and with each other in this co-located setting.

#### Method

We made hidden observations of a dozen of groups working in publicly accessible group learning spaces at our university, which each lasted about 15 minutes. Moreover, we conducted overt more in-depth observations with five groups, which lasted between 45 and 90 minutes each and which were followed by group interviews. Group 1 consisted of four male computer science students enrolled in the 7th semester. Group 2 was composed of four male first-semester students of chemistry. Group 3 were three male computer science students enrolled in the first semester. All three groups revised lecture topics for an exam. Group 4 consisted of three male economics students enrolled in their fourth semester. They prepared a term paper. Finally, Group 5, three male students enrolled in the fourth semester, met to solve the weekly exercises of a computer science lecture. Like in the other studies, participation was voluntary and no compensation was given.

Our study followed the research method of Grounded theory [Gla67]. This is a qualitative research method which aims at systematically generating a new theory from empirical data. In contrast to the more common method of formulating a hypothesis and then verifying or rejecting it with empirical data, a study which follows Grounded theory starts with collecting data *without* hypotheses. In Grounded theory, the unit of analysis is the incident. Already at the time of collecting empirical

data, the researcher marks key incidents with codes. In a first phase of data collection, every incident the researcher is aware of is coded in order to understand the problem domain. This results in a large number of codes, which are subsequently merged into a smaller number of more abstract concepts. When a core of concepts, which captures the essence of the problem domain, has been identified, new data is coded with these core concepts in mind without bothering with incidents of little relevance to the core. New data is now used to refine or modify the core concepts. In a further step, categories are formed from these concepts. These categories are finally the basis for creating a theory.

Following this method, we coded incidents during the observations and used observations and information from the interviews to refine and modify the coding schema. Amongst others, we coded which media are used, the physical positions of these media and how users interact with these media, for example reading, pointing, writing and moving. In subsequent steps, we derived the following theoretical concepts of co-located collaborative knowledge work with documents.

### **Functions of documents**

Three functions of documents became particularly evident:

1. *Reference*: Documents provide input about the contents to discuss. Reference documents include course scripts, books and web pages.
2. *Externalization*: Documents are used to externalize cognitive processes or the results of cognition. They illustrate own thoughts and document results of the learning process. This includes scribbling on scratch paper to foster shared understanding (temporary externalization) and condensing the contents in summaries written on empty sheets of papers or file cards (permanent externalization).
3. *Process structuring*: All but one group used documents as structuring scaffolds for their meeting. They followed the structure of the course script or of a list of exam questions provided by the instructor and discussed the topics in this sequential order.

### **Different paper and digital documents form one interconnected document space**

The meetings comprised a multitude of different document types. All groups used paper notebooks or empty sheets of paper and other paper documents (printed scripts, books, file cards or other documents). They moreover used physical tools for organizing and indexing paper documents, such as folders, ring binders and page marker stickers. About half of the groups also used electronic documents displayed on one up to three laptops. Laptops were mostly used for accessing lecture slides. Less frequently, users looked up information on web pages (above all Wikipedia) or



Figure 2.12: Media use in a typical learning group meeting

programmed code snippets in Eclipse. Most groups covered the available table surface to a high degree with documents. Figure 2.12 depicts a typical learning group meeting.

Although, at first sight, this multitude of documents may appear disparate and unclear, it forms one document space, in which the different documents are tightly interwoven, each having a particular function. Most tasks we observed include working with several documents at a time or shortly consecutively. The task of understanding a particular problem can for example include reading a slide of the course script on the screen of a laptop, making a sketch on a scratch paper at the same time and then formulating a summary on a new sheet of a paper notebook.

### **Interacting with people goes in hand with interacting with documents**

We observed both groups in which participants were constantly engaged in discussions and groups in which the participants worked rather individually on personal documents with short discussions coming up from time to time. A common finding in both extremes (and the levels in-between) is that almost all participants constantly interacted with documents all the time. During discussion, the most frequent activity was pointing to one's own documents or to documents of other group members (in the latter case mostly to documents of the person directly addressed when speaking). Another frequent activity consisted in moving documents on the table to allow shared working on a personal document or to retrieve a document to individually work on it. The joint or alternate writing in documents often occurred as well. The interaction with documents therefore seems to be a good basis for capturing the essence of a meeting.

### **Importance of the spatial arrangements**

We observed that the spatial arrangements of documents relates to ownership and to their current functions and importance. Documents that several people work with typically have another position than non-collaborative documents. The former ones usually are positioned between two persons or in the middle region of the table, at least during active collaboration. Different document functions became particularly clear with one group, which heavily relied on scratch paper. This was situated in front of one group member and collaboratively used. Each time, the current topic was understood, another group member wrote a synopsis and the used sheets of scratch paper were moved towards the border of the table. This way, they did not take up valuable space but still could be grasped quickly. Similarly, documents that are less frequently used were positioned at the outer zones of the table and moved to the center when needed. This demonstrates that group members implicitly partition the space available on the table surface into zones of different priority (cf. the distinction of hot, warm and cold zones of [SH03]).

### **Clear ownership of documents**

We found that learning groups clearly distinguish between personal and shared documents. Although personal documents may become a temporary focus of collaborative work (e.g. several persons reading and pointing to the same document), our observations show that their personal status is maintained: We never observed an initially personal document become a shared one during a meeting. Moreover, persons usually do not write in documents considered to be owned by other group members. Several groups used explicit shared documents which belong to all group members and which are usually copied after the meeting. These documents were scratch paper and handwritten summaries.

#### **2.3.4 Summary and Implications**

In a first questionnaire-based study, the majority of the participants indicated taking notes during the lecture and more than three quarters of them used only paper for taking notes. This is true despite the fact that 78.6 % of them owned a laptop. As most important factors for the choice of pen and paper, participants indicated the ease and flexibility of handwritten notes and the mobility of paper. The study further showed that participants judged important both to annotate printed handouts (e.g. lecture slides) and to take notes in empty areas or sheets of paper, as the latter allow creating one's own structure and noting own ideas more individually.

In a second study, we analyzed annotations on presentation slides made by students in a university seminar. A main finding was that the use of a keyboard seems to constrain the annotation process, as, in average, participants of the pen and paper group made more than twice as much annotations per participant and session than those of the laptop group. Another outcome relevant to our system design was that

the majority of the participants used specific classification schemes to mark document contents for later access.

A third ethnographic study of learning group meetings clearly indicated the amalgamation of printed and digital documents into one consistent document space. Typical spaces included paper notebooks, empty sheets of paper, printed scripts, books, file cards and additional physical tools for organizing and indexing documents. About half of the groups also used electronic documents displayed on up to three laptops. These multiple documents were tightly interwoven, each having a particular function. As these documents contribute to the same learning tasks, they form an interconnected document space that is used by several learners collaboratively. Participants leveraged the physical aspect of paper sheets and arranged them depending on the current working task.

These findings lead to implications for our research on several levels:

1. The finding that several learners collaboratively work with a single interconnected workspace lead us to applying an ecological perspective as our guiding theoretical framework. This will be discussed in the following section.
2. As a second implication of the results of the field studies – and of results of studies reported in the literature – we formulated requirements for the design of digital systems which support knowledge workers in learning with documents. These requirements will be presented in Section 2.5 at the end of this chapter.
3. Finally, we developed a theoretical model which identifies core interactions of digital pen and paper user interfaces and provides guidelines for the design of digital pen and paper user interfaces that are easy and intuitive to use. This framework will be discussed in Section 4.2.

## 2.4 An Ecological Perspective of Knowledge Work

Traditional cognitive approaches to human-computer interaction mainly focused on how an isolated user utilizes a computer system for performing an isolated task [RE94]. However, as we have seen, collaborative learning with documents is characterized by a joint and interconnected workspace of several learners. Individual activities occur in a complex network of users, documents, interleaving tasks and collaborative practices.

This lead us to adopting an ecological perspective of knowledge work as our guiding theoretical framework. This perspective surpasses the view of individual users and individual tools. It draws upon the concepts of Distributed Cognition [RE94, HHK00] and Information Ecologies [NO99]. Both theories commonly advocate a systemic view on the interplay of collaborating users, the use of physical and digital artifacts and the work practices. They argue that this perspective is the key for understanding and supporting knowledge work in a given work context. In the following, we will briefly present the central concepts of both theories.

### 2.4.1 Distributed Cognition

Distributed Cognition [RE94, HHK00] is a theoretical and methodological framework for the analysis of complex work processes which are socially distributed and which include the use of technological artifacts and tools. In addition to guiding the analysis, it aims at providing guidance for the design of novel systems which support these work processes.

The basic assumption of this theory is that cognition is not inherent to an individual. Instead, cognition occurs in the interactions between an individual and his or her environment (comprising other individuals and artifacts). Cognition is therefore embodied and situated within the work processes in which it occurs. Hence, the material world takes on a central rather than a peripheral role for cognition, as the work material becomes an element of the cognitive system itself.

Cognitive activities take place in a “functional system”: *a collection of individuals and artifacts and their relations to each other in a particular work practice*. Examples of functional systems include cockpits of airplanes, call centers and computer programmer teams. A cockpit, for instance, consists of pilots, artifacts, such as instruments, displays and printed flight manuals, and a set of practices, such as the precisely specified protocols for take-off or landing. In order to understand cognition in such a functional system, it is necessary to analyze the interaction of individuals, the ways they use artifacts and the influences of the environment. The focus is thereby on the question of how information is propagated across media, which means both internal media (the brain of an individual) and external media (e.g. a computer or paper media).

The navigation of a ship is a frequently cited example of distributed cognition. The activity of taking a bearing on a ship near the coast requires that one member of the navigation team finds landmarks on the shore, remembers their names and identifies the position of the ship with respect to these landmarks. This involves using specific tools, such as an alidade. He or she then reports these findings to another navigator in the pilothouse over a telephone circuit. This latter person then writes the information in a log book. A third navigator is responsible for plotting this information on a paper chart using various tools. The logger is either informed verbally or reads the information from the log book. This short example illustrates that “the outcomes that (matter) to the ship (are) not determined by the cognitive properties of any single navigator, but instead (are) the product of the interactions of several navigators with each other and with a complex suite of tools” [HHK00].

As Rogers et al. [RE94] state, Distributed Cognition informs the design of computing systems for collaborative work by analyzing how novel systems might fit into current work practices and in which aspects they might be disruptive. Ethnographic studies of functional units are a central method for these analyses. For example, at first glance it might appear reasonable to share documents electronically within an organization. Yet, an ethnographic analysis of the functional system might for example show that e-mail is disruptive, as personally handing a physical copy over to a co-worker might fulfill a communicative purpose other than just passing the

information contained within the document.

### 2.4.2 Information Ecologies

Although adopting a different theoretical approach, Nardi and O'Day [NO99] take on a similar viewpoint on knowledge work. Inspired from biological ecosystems, the authors introduce the idea that technology can be metaphorically seen as an ecology. They define an information ecology to be a *system of people, practices, values and technologies in a given local environment*. Their attention is on relationships involving tools, people and their practices.<sup>6</sup> Examples of information ecologies include libraries, self-service copy shops or intensive care units of hospitals.

Like a biological ecosystem, an information ecology is a complex system of parts (comprising people and tools) which have strong interrelations and dependencies. It contains a diversity of roles for the people and functions for tools. In a self-service copy shop, for example, there are such various tools as copy machines, computers, scanners, paper stock and scissors. If customers need help on how to use a machine, they can ask one another or get helped by the staff. Humans help other humans use technology.

Nardi and O'Day state that information ecologies are characterized by a continuous evolution. As novel technologies are integrated into current work practice, both technologies and practices are adapted and assimilated to fit to each other. For example, the introduction of word processing has dramatically changed the redaction and publication process of scientific literature. Nowadays, the scientific performs many of the tasks that were carried out by specialists, e.g. by typesetters, photographers and printers. Another example is the introduction of e-learning at universities. Distance and blended learning have partially changed the traditional practices of how we had learnt at universities before. Currently, this evolution of practices causes in turn an evolution of technology. Instead of purely relying on technology for distributing content to the learner (an approach which is often denoted as eLearning 1.0 and now considered somewhat old-fashioned), technology is developed further to support learners in creating and sharing their own content (eLearning 2.0 [Dow05]).

A further important aspect of information ecologies is what the authors call their "locality". The same technology (e.g. the same type of computer with the same hardware and software configuration) can be used very differently in different environments. Hence, the local participants define the identity and place of the technology. As a consequence, it is the task of the designers to provide useful functionality, but the local participants complete the job by integrating them into their practices in a way that makes sense for them.

---

<sup>6</sup>Note that this is very similar to the functional system of Distributed Cognition. People correspond to individuals, practices and values of a local environment correspond to a particular work practice and technology corresponds to artifacts.

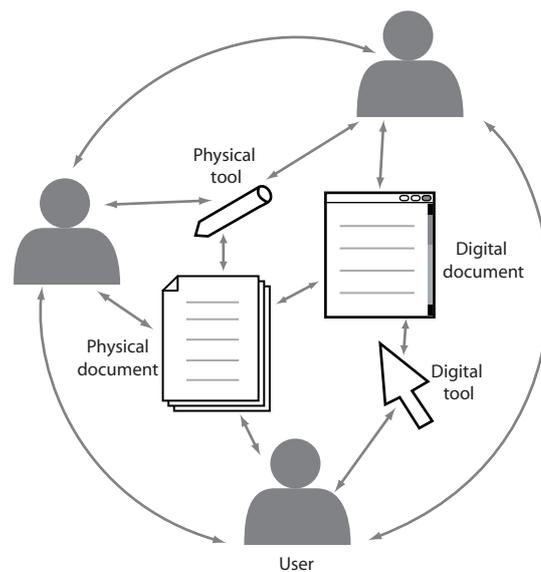


Figure 2.13: Main elements of an information ecology

### 2.4.3 Ecological Perspective

This brief review demonstrates that both theories, although different in their approach, are very similar in that they focus on the interrelations between actors, technology and given practices for understanding the use of technology in work settings. Based on both theories, we define the ecological perspective as follows:

**Definition 3** (Ecological Perspective of Knowledge Work). *An ecological perspective of knowledge work is the integrated, systemic analysis of the elements of a particular work setting. This consists of the users, of the physical and digital artifacts and tools, of the practices of using these artifacts and tools, of the relations between users as well of the relations between artifacts and tools.*

In our field of application, main elements of an information ecology are documents in physical or digital form, physical and digital tools that support working with these documents (e.g. pencils, rubbers, ring binders, computers, mice, screens and printers), users, practices as well as the relations between documents, between users, between documents and tools and between users and documents. These elements are depicted in Figure 2.13.

**Consequences for this thesis.** Adopting this perspective as our guiding theoretical framework has several consequences for our research:

- First, it underscores the *need for thorough field studies* in order to assess these interrelations in the field which shall be supported by novel systems. It is

crucial to analyze the main activities and practices, the roles of people and the use of artifacts.

- As a second consequence, our solutions are designed for *collaborative use*, since collaboration is an inherent part of knowledge work.
- Third, the viewpoint that cognition is embodied in the material world provides a *theoretical explanation for the fact that users prefer paper to computers* in many work settings. For instance, organizing paper documents by making piles on a desk embodies the cognitive activity of structuring a problem domain in the material world.
- Fourth, the distributed cognition approach, which focuses on how information is propagated across media, brought us to assess how information of several documents is related to each other. Amongst others, we found that this propagation often goes in hand with an abstraction (e.g. writing a summary based on existing documents). These relations are typically implicit and cannot be easily understood by others. We therefore designed *mechanisms for making these relations explicit*.
- Finally, we opted for designing *generic concepts* which can be “localized” by the end users to make them fit into their given work practice.

## 2.5 Requirements

Based on the findings of the field studies and on the ecological theoretical perspective, we infer the following requirements for an interaction design which supports collaborative learning with documents.

### R1. Support of both printed and digital documents

We have seen that both printed documents and digital documents are typical artifacts in information ecologies in university learning. As we have seen, the main reason for the persistence of paper is that even state-of-the-art digital technologies, such as Tablet PCs or E-Book readers, cannot emulate the most important advantages of paper. Although the user can print digital documents onto paper, the reverse direction is not supported. Changes made to printed documents must be manually (re-)integrated into a digital version of the document. This is time-consuming and moreover runs the risk of two inconsistent document versions.

The design should therefore *support the digital interaction with printed documents*. This combines the flexibility and the mobility of paper with the capabilities of computing. In addition, it shall support similar interactions with digital documents. This enables users to choose the adequate medium for a given task. On a technical level, this integration of paper and digital user interfaces can be realized by a class of interfaces that we call Pen-and-Paper User Interfaces. We define these as follows:

**Definition 4** (Pen-and-Paper User Interface (PPUI)). *Pen-and-Paper User Interfaces (PPUIs) integrate interactions on real paper with digital support. One or several paper sheets contain printed user interface elements. The main interaction device is a digital pen whose position on the paper sheet is digitally tracked. This provides for entering digital data by writing and drawing on the paper sheet. Moreover, the user may enter additional interactional information (e.g. issuing a command by performing a pen “click” on a virtual button which is printed on paper). The system provides digital feedback via the pen itself (built-in display, LEDs, audio or haptic feedback), by printing an updated version of the paper sheet(s) and/or by providing audio or visual feedback using nearby devices, such as PDAs, displays or beamers.*

The specifics of PPUIs lead to the following requirements:

**R1a. Similar interactions in the physical and in the digital worlds.** In order to provide a close integration of physical and digital documents, the system shall offer similar interactions for printed and digital documents. This requires

1. a *similar representation of documents* regardless if they are printed or displayed on a screen
2. that the *same interaction device* should apply to documents in both representations
3. that the *interactions* to be performed with this device are *as similar as possible in both worlds*.

**R1b. Reliable interactions even in the absence of digital feedback.** A problem with PPUIs is the restricted feedback loop. As it is not viable to re-print the paper user interface for each minor update of the digital system’s state, the designer must resort to feedback media other than the input medium. Current digital pens can provide only very restricted feedback. Therefore, external devices, such as nearby screens, are typically used for this purpose. Yet, the key affordance of pen and paper is its mobility and the fact that no other devices or tools are necessary during mobile use. To retain this key affordance, interactions in a PPUI shall be reliable and clear even if no immediate digital feedback can be provided.

## **R2. Support of annotating, linking and tagging collections of documents**

Based on the results of the field studies, we identify three main conceptual activities to be supported in working with existing documents<sup>7</sup>.

---

<sup>7</sup>Learning with existing documents stands in contrast to producing new documents, such as composing an article. While learning with existing documents might include producing documents (for instance a summary or a glossary), producing documents is not the primary focus of this activity.

**Annotating and notetaking.** First, the design shall enable users to make *handwritten annotations* on existing documents and to *take notes* on empty pages. We have seen that this is a very important activity which constantly occurs in document-based learning. Psychological research shows that notetaking plays an important role in learning processes and has been proven to be a factor positively related to students' academic achievement [Kie87, POK05]. This supportive effect encompasses both the processes of recording notes and annotations (encoding function) and reviewing notes later on (storage function) [Kie89]. Accounting for the highly individual annotation style encountered in the field studies, the paper-based user interface should impose as little constraints as possible on the flexible interaction with paper documents.

**Relating several documents.** The design shall further support users in *working with several documents at a time*. We have seen that typical information ecologies include several documents between which the user frequently jumps back and forth, because the different documents are related in manifold ways. These relations often span the paper/digital boundary. Today, they typically are not expressed in an explicit manner, but the user is only implicitly aware of them. This pure implicit management has several disadvantages. When not working with a document collection for some time, the user might forget some or all implicit relations. This makes it difficult to understand and efficiently use the complex web of information in a document collection later on. Moreover, the pure implicit knowledge about relations is hard to share with co-workers and the invaluable information about the structure of a document collection can not be leveraged by digital systems.

**Hyperlinking.** For these reasons, the system shall enable the user to *express these implicit relations in an explicit manner by creating hyperlinks* between entire documents and passages of documents, regardless if they are available in a printed or in a digital representation. Similar to Vannevar Bush's early vision [Bus45], the knowledge worker will be able to create a personal web of information. In contrast to Bush, this closely connects physical with digital media.

**Tagging.** Third, the design shall offer the possibility to *tag printed and digital documents*. The field studies have shown that users dispose of a wide range of traditional techniques for structuring documents and collections of documents. This includes semantic classification schemes with user-defined symbols, practices of efficiently using spatial arrangements for expressing relatedness and priority of documents and a whole bunch of paper-based indexing tools, such as bookmarks, page marker stickers, folders and ring binders. In addition to structuring documents with hyperlinks, tagging enables learners to further integrate new knowledge with existing knowledge by relating information and abstracting it to concepts. This generates structural knowledge of how concepts of a domain are interrelated. The knowledge of these relations and the ability to explain them is existential for higher-order

procedural knowledge [JBY93].

In practice, annotating, linking and tagging go in hand and *complement each other*. When reading a document, the knowledge worker might for example make handwritten annotations to highlight important passages and to add some clarifying details. Moreover, he or she might add a reference to another document which covers a particular aspect in more detail. When reading this other document, he or she makes annotations again before tagging this other document as important and finally going back to the first document.

**Genericness.** Annotating, linking and tagging are *generic activities*, each of them serving a large number of possible tasks. For instance, free-form annotations support tasks such as taking notes during lectures, structuring documents, revising documents, creating summaries or making excerpts. In each of these tasks, the user creates other types and forms of annotations at other document positions. These various practices are possible because annotations can be of any shape and can be made at arbitrary positions within documents. Another example of generic activities are tags. These can be used for such different tasks as conceptually structuring a given domain, defining priorities and to do items or selecting passages that a co-worker should read by tagging them with the name of this person. This generic character of the activities implies that the system can be used in a wide variety of knowledge work settings. In our application scenario of learning at universities, this includes taking notes and making annotations during courses, reviewing own notes and shared notes of other learners, preparing for exams in learning group meetings, excerpting documents, searching and integrating literature for preparing an article or a term paper and even giving presentations by controlling the slide actually being presented using a printout.

### **R3. Support of co-located and remote collaboration**

The design shall support the *various forms of collaboration* that exist in university learning. This includes co-located collaboration, for example in learning group meetings, and remote collaboration.

Collaborating with physical documents is characterized by several inconveniences. Using printed documents is very powerful during co-located interaction, because their physical nature affords flexible multi-user interaction in co-presence. For example, it is very easy to jointly write on a document and documents can be flexibly moved and spatially organized to structure them or to (re-)attribute specific documents to specific persons. However, after a co-located work phase, the resulting physical documents are hard to share with all co-workers. The same is true for remote collaboration, where it is tedious to share information contained in printed documents which is not digitally available. In this context, the automatic digitization of interactions with physical documents offers a benefit that shall be leveraged to support collaboration.

#### R4. Symbiotic integration into current information ecologies

The ultimate goal of a new interaction design is that users are able to symbiotically integrate the new practices into their existing information ecologies. This comprises three main aspects:

**R4a. Interactions that produce little extraneous cognitive load.** Knowledge work and particularly learning with documents is typically characterized by a high degree of intrinsic cognitive load [CS91], as the knowledge worker aims at acquiring or extending existing knowledge or generating new knowledge. For instance, reading and understanding complicated subjects matters, which are detailed in a document, or composing new documents are cognitively demanding tasks.

It might be by no means a simple coincidence that traditional tools and interactions for working with documents are characterized by a high degree of simplicity and intuitiveness and therefore produce little extraneous cognitive load. As we have seen in Section 2.2 and in Section 2.3, this provides for interactions that are performed in an implicit manner, that are interwoven with other tasks and that remain behind what is really important – the actual knowledge work task to be supported.

Like this traditional practice, novel interactions shall produce little extraneous load in order to be easily integrated into the information ecology.

**R4b. Rich and flexible interactions.** A second challenge is connected to the implicit way we use traditional pen and paper. The literature and our field studies have shown that practices of working with printed documents are highly individual. For instance, this comprises the way how documents are marked up during reading and the way how collections documents are organized.

It is therefore crucial that digital support accounts for the flexibility of traditional practices. From this it follows that users shall have a *choice amongst a set of interactions* to select the one that fits best the individual working style and current context. Moreover, individual interactions shall be *adaptable to the user's preferences and habits*.

**R4c. Structured access to all information.** A digital system which integrates interaction with printed and digital documents and collaboration, shall *provide an integrated and structured access to the information ecology*. This stands in contrast to current digital support, which typically focuses on single documents instead of the entire ecology.

The system disposes of many information, which integrates and relates many documents and several users: As the system collects user generated content on documents, it disposes of the annotations, it has information about the relations between documents and it has knowledge about which tags apply to which documents or passages of documents. In addition, digital pen technology provides temporal information about the interactions and enables to attribute activities to individual users.

All this information of various dimensions (documents, user generated content, activities, users, time) shall be automatically integrated and made available in novel visualizations. These shall provide a structured overview and access to the information ecology.

## 2.6 Conclusions

In this chapter, we approached the field of knowledge work from a socio-scientific perspective focusing on learning at universities. This includes a variety of practices ranging from guided to highly self-autonomous learning and from individual to collaborative learning.<sup>8</sup> In this chapter, we examined how documents are used and how users collaborate in knowledge work and learning. Our analysis is based on findings reported in the socio-scientific literature and on the outcomes of three field studies, which were conducted as part of this thesis.

**Paper as a key information medium.** The analysis shows that paper is still a key information medium in knowledge work. This has several reasons. First, paper documents can be very easily navigated with both hands. Second, they can be easily and very flexibly annotated with a pen. Our questionnaire-based study with more than 400 students showed that more than three quarters of them use only pen and paper for annotating lecture handouts, even though the large majority owns a laptop. An analysis of more than thousand annotations made by students in a university seminar found that the use of a keyboard seems to constrain the annotation process. In average, participants using pen and paper made more than twice as much annotations than participants using an annotation tool for textual annotations on a laptop. A third quality of paper is the fact that paper embodies information. This provides for organizing and structuring information by arranging documents the physical space. Moreover, paper affords to seamlessly interweave multiple activities at a time, for example reading with writing. This heavily draws on the intuitive and two-handed interaction as well as on the concurrent use of many display surfaces. Finally, paper is highly mobile and supports various forms of implicit and explicit collaboration as well as mutual awareness in co-located settings. Due to these affordances, paper efficiently supports reading documents, processes of planning and thinking as well as collaboration and organizational communication – all highly relevant processes for knowledge work.

**Integration of paper and digital media.** As a matter of course, these advantages of paper do not imply that knowledge workers exclusively work with paper. Digital documents are also key parts in today's knowledge work. In a third field study

---

<sup>8</sup>Despite the focus on the learning domain, most results of this thesis can be transferred to other knowledge work settings, for example to scientific research or to industrial planning, research and development.

of learning group meetings, we found that printed and digital documents are frequently amalgamated into one interconnected document space that is collaboratively used by several learners.

These findings render obvious that technology for knowledge workers should not exclusively support digital media. Instead, it should provide for flexible transitions from paper to digital media and vice versa. A promising type of user interface for such systems are Pen-and-Paper User Interfaces (PPUIs). By writing with a digital pen on real paper, users combine the advantages of paper with the powerful capabilities of computing.

**Ecological perspective.** Moreover, the findings of our studies brought us to choosing an ecological perspective as the guiding theoretical framework for knowledge work with documents. The ecological perspective advocates an integral view on the interplay of collaborating users, the use of physical and digital artifacts and the practices in a given work setting. This is the key for understanding and supporting knowledge work in a given work context, but has not been applied to Pen-and-Paper User Interfaces beforehand.

**Requirements.** From this ecological perspective, we can directly derive several requirements for a system that supports knowledge work with documents. These are summarized in Table 2.2. The system shall tightly integrate printed and digital documents and offer similar interactions in both worlds. The design shall support the various forms of collaboration that exist in knowledge work including co-located collaboration and remote asynchronous collaboration. Moreover, the design shall provide means for enabling a structured and integrated access to the information of the ecology. This includes information contained within printed and digital documents and their relations as well as information about users and their activities. Further, novel interactions are to be symbiotically integrated into current information ecologies. For this reason, they shall produce little extraneous cognitive load and be adaptable to the user's habits and preferences.

Based on our field studies, we identified three main conceptual activities around documents that complement each other, are generic and can be used for a wide variety of tasks. Users shall be enabled to extend printed and digital documents by annotations, to relate documents with hyperlinks and to abstract to higher-level concepts with tags. Annotating is an important activity during active reading and thinking. Moreover, by abstracting and establishing relationships between concepts, learners build structural knowledge, which facilitates recall and comprehension and is essential to problem solving.

In this chapter, we have acquired socio-scientific background knowledge of the problem domain, chosen a theoretical foundation for our work and derived requirements for interactive systems that support knowledge work with documents. On the way towards designing system support, it is an important next step to review the current state of the art of technologies and systems which aim at integrating paper with digital media. Therefore, in the next chapter, we survey prior research on

| <b>Requirement</b> |   |
|--------------------|---|
| R1                 | Support of both printed and digital documents                       |
| R1a                | Similar interactions in the physical and in the digital worlds      |
| R1b                | Reliable interactions without digital feedback                      |
| R2                 | Supporting annotating, linking and tagging collections of documents |
| R3                 | Supporting co-located and remote collaboration                      |
| R4                 | Symbiotic integration into current information ecologies            |
| R4a                | Interactions that produce little extraneous cognitive load          |
| R4b                | Rich and flexible interactions                                      |
| R4c                | Structured access to all information                                |

Table 2.2: Overview of the requirements identified in this chapter

Pen-and-Paper User Interfaces focusing on the requirements that we have identified in this chapter.



# 3

## State of the Art

In the previous chapter, we have approached the problem domain from a socio-scientific perspective. Based on results from the social sciences literature and on own user studies, we identified affordances of paper as well as practices of using paper. We finally derived requirements for systems that offer support for knowledge work with documents.

This chapter reviews the state of the art of previous work that addresses the problems identified in the previous chapter. As our first requirement is the tight integration of paper and computers, this chapter focuses on Pen-and-Paper User Interfaces (PPUIs). We start in Section 3.1 with a theoretical view on interaction with digitally augmented paper and summarize prior research on theoretical models. We show that while these model the relations between physical objects and the digital information they represent, these do not sufficiently describe interactions with pen and paper. Next, in Section 3.2, we review technologies that can be used for identifying physical paper sheets, for digitizing the contents of paper documents and for digitally tracking the interaction with paper. This section shows that there is a wide variety of sophisticated technologies available which offer support for these tasks. In Section 3.3, we discuss the state of the art of applications that integrate printed with digital documents. Thereby, our focus is on the considerable body of research on interfaces that support annotation, linking and tagging of documents. This survey shows that prior work has significant shortcomings concerning the support of collaboration with printed documents, the perceptual coupling of printed and digital media and the ease of integration into existing ecologies.

This chapter was written with the goal not only to justify the research of this thesis, but also to provide a more general state of the art survey. For this reason, related work is first described in a neutral manner before we provide a comprehensive critical discussion with respect to our requirements in Section 3.4.

### 3.1 Theoretical Models of Pen-and-Paper User Interfaces

Pen-and-Paper User Interfaces (PPUIs) rely on different interaction metaphors and techniques than the well-known class of Graphical User Interfaces (GUIs). It is therefore an important research question how interaction with digitally augmented paper can be analyzed and classified with a theoretical point of view. Theory both helps in structuring the design space, in analyzing and categorizing existing applications and in guiding the design of novel PPUIs.

Almost all research on PPUIs focuses on systems and not on theory. In contrast, in the field of Tangible User Interfaces (TUIs) – of which PPUIs are a subclass – a growing number of publications address concepts and theory. These concepts of TUIs can be used to theoretically describe partial aspects of PPUIs. We therefore start by briefly reviewing this research. However, as this work focuses on general TUIs, it offers little guidance for the specifics of interacting with a pen in a paper environment. For this reason, in the following we also discuss the theoretical work which has been conducted on PPUIs and identify open issues.

#### 3.1.1 Models of Tangible User Interfaces

Theoretical frameworks of Tangible User Interfaces have made important contributions for the theoretical structuring of the domain by defining main components and conceptual terms and by categorizing systems. We can distinguish the following main dimensions which are addressed by these frameworks:

1. The coupling between physical and digital input and output.
2. The conceptual and perceptual similarities between physical objects and their digital counterparts.
3. The interaction with tangible objects in the physical space.

**Coupling between physical and digital input and output.** Ullmer and Ishii [UI01] extend the well-known Model-View-Controller paradigm [KP88] to an interaction model for combined physical and digital interfaces. This describes elements for physical and digital input and output. It carries over the ‘model’ and ‘control’ elements while the ‘view’ element is divided into two subclasses. This accounts for the fact that most tangible interfaces represent the model’s state by a combination of physical and digital information. The first subclass of the ‘view’ element are tangible objects which represent (parts of) the state of the digital model. These objects can also act as a physical control for the digital model. Second, a digital representation can provide further information on the system state, e.g. on dynamic information which is hard to express by physical objects. For example, this digital representation can be provided on a nearby screen, on a speaker or by projecting a digital image onto the tangible objects.

**Conceptual and perceptual similarities.** The taxonomy of Fishkin [Fis04] includes two dimensions. The first dimension, which is called ‘metaphor’, focuses on the perceptual similarity between physical objects and their digital counterparts. This similarity concerns both the physical shape/look/sound of the object and the actions that are performed with this object. The second axis (‘embodiment’) corresponds to the spatial distance between the tangible object which is used for input by the user and the device that provides system feedback. Fishkin distinguishes four levels of embodiment ranging from feedback which is provided on the input tangible to feedback which is provided on a distant device. Fishkin states that, in order to generate the impression of computation being embodied within the tangible objects, the tangible input device should also be the output device. However, in other situations, a spatially more distant system feedback might be more appropriate. Besides this spatial offset, Fishkin’s taxonomy does not address temporal offsets. These occur in scenarios where physical interaction cannot be tracked in real-time by the digital system. An example is the situation in which a digital pen temporally buffers data when used in a mobile setting before being synchronized with a computer later on. The distance between input and output is also addressed by the framework of Koleva et al. [KBNR03], which denotes this as the degree of coherence.

**Interaction with tangible objects.** Other frameworks analyze the concrete interactions that the user performs with tangible interfaces. Interaction in most TUIs is centered on moving and arranging physical objects. For instance, the seminal URP system [UI99b] enables urban planners to modify a digital model of urban buildings by moving physical models of these buildings. With other systems, such as the Marble Answering Machine [IU97] or MediaBlocks [UI99a], the user accesses and modifies digital information by moving and arranging objects that act as physical handles for this information. Correspondingly, theoretical approaches to interaction within TUIs conceptualize interactions as changing the location or orientation of objects.

The framework of Ullmer and Ishii [UI01] classifies TUIs by the way in which they combine multiple tangible objects. The TAC paradigm [SLCGJ04] states that it is the physical constraints that define which interactions are possible (and not possible) with tangible objects in a TUI. Again, interactions are conceptualized as displacements and compositions of tangible objects. These concepts do not account for other types of interactions that alter the tangible objects themselves rather than displacing them. Ishii and Ullmer [IU97] transfer a set of GUI elements to TUIs (such as windows, icons and handles) suggesting generic physical instantiations of these elements. The focus is again on interaction as displacements, rotations and compositions of objects. Interactions with an individual object have meaning only with respect to other objects or to a reference frame.

All these theoretical frameworks do not account for the collaborative use of TUIs by multiple users. The framework of Hornecker and Buur [HB06] briefly discusses co-located use of TUIs by pointing out that TUIs offer multiple points of interaction,

which provides for spatially distributed control. Nevertheless, the collaborative dimension has not been extensively analyzed in TUI models.

### 3.1.2 Models of Pen-and-Paper User Interfaces

In contrast to general TUIs, research on Pen-and-Paper User Interfaces (PPUIs) almost exclusively focused on developing new systems. In the following we will review research that takes on a more theoretical point of view.

Guimbretière introduces a lifecycle model of transformations between paper and digital documents [Gui03]. Although not presented as such, it is an important counterpart to Ullmer and Ishii's interaction model [UI01], which was mentioned above. Ullmer and Ishii model how *different* physical and digital representations are used simultaneously in a complementing manner (e.g. a digital projection overlaying physical building models). In contrast, Guimbretière models how the *same* document can be accessed in equivalent physical and digital representations. Depending on the situation, the user chooses the representations that best fits her needs. For example, in a mobile setting, the user might prefer reading and annotating a printout of the document, while she prefers working with a digital representation for sharing it with co-workers. Offering the user both a physical and a digital representation of the same information and letting the user choose between both of them is a dimension which is to the best of our knowledge not considered in the research on general TUIs.

Yeh et al. [YPK08] define a design space of paper interactions and present a toolkit for the rapid development of PPUIs. The toolkit offers generic elements for printed user interfaces, including input fields for handwritings and sketches, buttons and check boxes. However, the underlying (implicit) interaction model focuses on interactions with single sheets of paper only leaving aside the important dimension of physical arrangements of pages and of interactions that span multiple pages (such as Pick-and-Drop [Rek97] and pen-based stitching gestures [HRG<sup>+</sup>04]).

Holman et al. [HVA<sup>+</sup>05] discuss how we might interact with documents on multiple digital paper displays, which are light, flat and malleable like paper. They introduce interaction primitives amongst others for activating and printing documents, for copy&paste and for scrolling within documents. These interactions rely on physical manipulations of paper displays, such as picking them up, collocating, flipping and stapling them, and not on pen-based interaction.

Finally, the iServer and iPaper framework [NSW06a] presents an extensive generic model for links between physical and digital documents. However, it does not cover interaction techniques.

### 3.1.3 Discussion

Existing theory offers appropriate models for analyzing the interplay between physical and digital representations and for assessing the relations between physical objects and the digital information they represent. However, on the level of interac-

tions, existing models do not sufficiently describe Pen-and-Paper User Interfaces. As discussed above, models of Tangible User Interfaces consider only the interactions of displacing, rotating and arranging physical objects while individual objects (e.g. shape, texture, content) are considered static. In contrast, interaction with pen and paper comprises writing with a pen. This implies that the tangible objects, i.e. the paper documents, themselves are permanently altered.

An interaction model of Pen-and-Paper User Interfaces should identify generic interaction primitives that are performed with pen and paper. A first step towards these interaction primitives is the work by Yeh et al. [YPK08]. However, this models only interactions with individual sheets of paper and not interactions that span multiple pages. As discussed in Section 2.2, main advantages of paper are precisely these multi-page interactions. In the next chapter, we will present a generic interaction model that accounts for this dimension.

## 3.2 Technology

Pen-and-Paper User Interfaces (PPUIs) require technologies for capturing both the contents of paper documents and the interactions the user performs with these documents. In this section, we review these core technologies that are at the bottom of PPUIs.

A first class of technology provides for uniquely identifying physical pages and their position. A second class captures the contents which are contained on sheets of paper. Finally, there exist approaches that capture the live interactions that a user performs on paper, such as pointing or writing. In the following, we discuss each of these approaches in turn. All these technologies aim at digitizing *input* from paper. Another direction of research is using paper as a digital *output* channel (e.g. [CHBM07, CP08]). This is not covered by this survey.

### 3.2.1 Page Identification and Location Tracking

Many settings that integrate paper with computing require that a physical sheet of paper can be uniquely identified. For instance, a physical paper card could be used as a physical token to access a specific digital object. Hence, the system must identify the card. A second example is a system that tracks the location of a physical documents in order to help the user in finding them later on. Two main approaches can be distinguished: vision-based and RFID-based approaches.

**Vision-based approaches.** Vision-based approaches typically rely on barcodes. A barcode encodes a unique identifier in an optical machine-readable representation. It is captured by a 2D camera or by a laser scanner. The oldest form of barcodes are linear (one-dimensional) barcodes. These encode the identifier using a pattern of bars of variable width. A very frequent linear encoding is EAN 13 [ISOa], which is



Figure 3.1: Examples of linear (left) and two-dimensional (right) barcodes.

used worldwide to identify products at cashpoints. Figure 3.1 (left) shows an example of this barcode. 2D barcodes encode data in a two-dimensional grid of points. An example is the DataMatrix [ISOb] code (Fig. 3.1 right), which is for example used for electronic stamps by the German postal service (Deutsche Post). Three-dimensional barcodes encode a third dimension by varying visual properties of the points, such as color or brightness.

In addition to *identifying* objects, vision-based tracking can also be used for tracking the *location* of objects. A camera takes two-dimensional images of the documents which contain two-dimensional barcodes. The four corners of a barcode enable to extract the barcode's relative position with respect to the camera with 6 degrees of freedom. Several toolkits exist that offer out-of-the-box support for application developers. The currently most widespread toolkit is ARTag [Fia04]. This supports up to 1024 barcode markers which have been optimized for fast and reliable detection. Another influential toolkit is ARToolkit [PKB05].

Main advantages of using barcodes for identifying paper documents are that barcodes are a relatively cheap tracking solution, as they can be printed with the document. However, this approach requires that the barcode is in the line-of-sight of the camera, possibly restricting natural interactions. Moreover, the artwork of the document is changed because the barcode is visible to the human eye. Recent research demonstrated a new technology of invisible visual markers [KNF09]. This relies on polar filters. However, these markers cannot be printed with the document but must be applied in a separate processing step.

**RFID-based approaches.** The second approach, RFID-based tracking, avoids the problems of barcodes. Radio-Frequency Identification (RFID) uses tags that are embedded into or applied to a physical object. The tag consists of an integrated circuit that stores a unique ID and manages the communication with an external reading device. Moreover, it includes an antenna for receiving and transmitting signals. If the tag is passive, it does not include an own battery but receives energy from the



Figure 3.2: The desktop scanner HP ScanJet N9210 (© Hewlett-Packard)

reading device via an electromagnetic field. When being within the range of a reading device, the RFID tag transmits its unique ID. No direct line of sight between the tag and the reading device is necessary. However, in contrast to barcode tracking, RFID tags produce additional costs. Applying them to a paper page requires an additional processing step beyond printing. Moreover, while the technology allows to track if a tag is within the range of a reader, it does not allow to further detect its precise location.

Besides these two main approaches, some other systems use individual solutions for identifying objects. Similarly to RFID, integrated circuits storing a unique identifier are applied to physical objects. The communication to a reading device is made via a wired connection (e.g. [SJH<sup>+</sup>08, JM09]) or a wireless connection (e.g. [nPSN00]).

### 3.2.2 Digitizing the Contents of Paper Documents

Visual scanning of sheets of paper is a well-established approach for digitizing the contents of paper documents. This captures an image of the page contents at a given point in time. If the image contains text, optical character recognition (OCR) technology [PS00] can be used to convert the graphical marks into a machine-readable symbolic representation.

Desktop scanners are well-suited for scanning large numbers of pages and offer a high resolution. For example, the commercially available HP Scanjet N9210 [HPS] (Fig. 3.2) digitizes the contents of up to 50 pages per minute with a resolution of 600 dots per inch. In contrast to desktop scanners, handheld scanners are small and light and can be used in mobile settings. A first class of handheld scanners has the scanning unit in the tip of a pen-like device. The interaction for scanning resembles to using a pen on a paper document. By moving the pen along the lines of text, the document is scanned little and little. This approach is well-suited for scanning individual words or short passages but too slow for scanning entire pages. An example is the Wizcom InfoScan2 Elite pen [Wiz] (Fig. 3.3 left). The pen is able to scan with a speed of about 15 cm/s and a resolution of 400 dpi and includes optical character recognition. This enables the pen to read loud the scanned text



Figure 3.3: Left: Wizcom InfoScan2 Elite Pen (©Wizcom Ltd.). Right: Planon DocuPen RC 805 (©Planon Ltd.)

using voice synthetization, to translate scanned text and to provide definitions for scanned words. It includes a display for accessing scanned documents. The user can also transfer scanned data to a computer or a PDA via an infrared port or via USB. A second class of handheld scanners is not used vertically like a pen, but lies flat on the document, much like a ruler. The scanning module therefore has a larger width, which significantly speeds up the scanning processes. This enables scanning an entire page in 4 to 8 seconds. An example is the Planon DocuPen RC 805 [Pla] (Fig. 3.3 right).

The advantages of desktop and handheld scanners are that they can scan all types of paper documents without additional provisions made to the documents. However, as a still picture is taken at one given point in time, it is not possible to continuously track changes of the document. Moreover, all visual contents are digitized in one single layer. For example, it is therefore a complex task to separate the conceptual layer of handwritten annotations and sketches from the underlying printed document.

A further approach captures the documents contents with one or several cameras which are mounted above the user's desk or in front of an interactive wall. This is well-suited for interactive systems, as the contents can be continuously tracked. We will discuss this approach in more detail below.

### 3.2.3 Tracking User Interaction on Paper

We have discussed technological means for identifying physical pages and tracking their locations. We have also seen how the contents of a paper document can be digitized. A further important field for bridging paper and computers is to track user activity on physical pages. This enables interactive systems in which the user can interact with pages, for example by tapping on hot-spot areas or by writing on it.

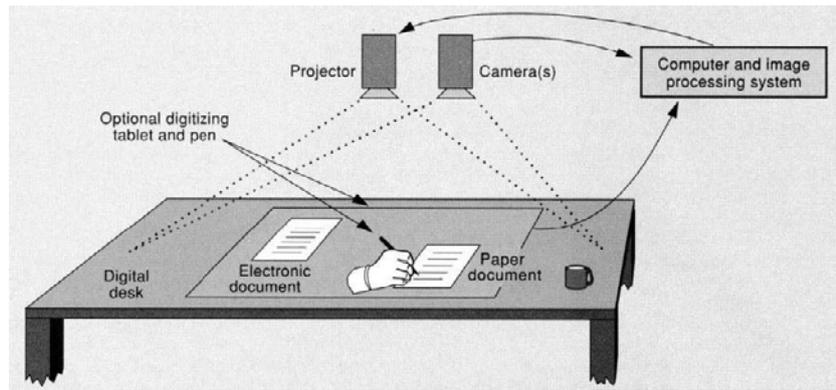


Figure 3.4: Schematic view of the Digital Desk [Wel93]

### Hand and Finger Input

Above, we have already presented vision-based approaches for page identification, location tracking and content capturing. This technology can also be used for tracking live user interaction on pages. A seminal system is Wellner's Digital Desk [Wel93] (Fig. 3.4). A camera which is mounted above the desk captures the position and the contents of printed documents and tracks gestures that the user performs with her fingers or with a pen. A beamer projects additional digital documents onto the same surface. This creates very seamless integration of physical and digital contents. The example applications of the Digital Desk include a calculator application. Instead of manually entering numbers, the user can copy numbers from printed documents by pointing on these. For this purpose, optical gesture recognition and optical character recognition is used. Several successor systems of the Digital Desk further improved hand tracking and gesture recognition [KK98, HVA<sup>+</sup>05] and designed further applications [RSW<sup>+</sup>97].

A drawback of camera-based scanning is that it significantly restricts the mobility of paper. Typically, the camera is mounted at a fix position. While there exist mobile solutions, such as the Docklamp [KDL07] or Sixth Sense [MMC09], these are still rather large and heavy and their use is not comparable to the flexibility of traditional pen and paper. Moreover, in any case, the paper documents must be kept in a small area in front of the camera.

### Pen Input

One specific form of interaction which is central for paper documents is pen-based input, such as handwritings and sketches. There exist capturing technologies that focus on this specific content.

Digital pens can be used in a manner which is similar to using traditional pen and paper. They leave visible ink traces on the paper sheet but additionally track their position. This generates digital ink data, a digital two-dimensional representation of



Figure 3.5: Left: Pegasus Tablet Notetaker (© Pegasus Ltd.). Right: Wacom Graphire4 graphics tablet (© Wacom)

the pen traces which moreover includes temporal information. Some technologies also track the force with which the pen tip is pressed onto the paper sheet.

Technology for capturing pen input on real paper should offer high tracking performance while restricting the natural interaction as little as possible. We distinguish approaches that detect a relative pen position with respect to an external reading device from approaches that detect absolute positions on the paper pages.

**Relative position tracking.** A first class of approaches tracks the relative position of a digital pen on a fixed surface with respect to a separate tracking device. The most widespread solution within this class relies on ultrasonic triangulation. The digital pen continuously emits an ultrasonic signal which is not audible by humans. A tracking device is attached to the paper sheet(s). It has two or more reference points that detect the pen distance using the signal. Based on ultrasonic position detection, the position of the pen with respect to these reference points can be calculated. The resolution of this technology is high enough for capturing handwriting. As the tracking does not depend on the material the pen is used upon and scales to large surfaces, it can be used on arbitrary surfaces, such as tables, augmented walls or whiteboards. However, the position of the pen is tracked in relation to the external device and not in relation to the paper sheet. Therefore, the user must manually calibrate the position of the sheet and once calibrated, a page must not be moved. Moreover, this approach is not able to detect which sheet the user is writing on. For this reason, the user must manually indicate when he is changing the page. This makes this approach hard to use in settings where users do not write on one single page but deal with many pages. Ultrasonic tracking is utilized in commercial solutions that mainly focus on private end-users, for example the Pegasus Tablet NoteTaker [Peg] (Fig. 3.5 left). This has a resolution of 100 dpi.



Figure 3.6: Left: Magnified view of a printed document containing the Anoto pattern. Right: Schematic view of a 6x6 grid of points (© Anoto AB)

Another type of external tracking device are graphics tablets. This is a flat surface which generates a magnetic field. Using induction, the position of a specific stylus can be detected on this surface. Figure 3.5 (right) shows a Wacom Graphire4 [Wac] graphics tablet. In principle, graphics tablets do not aim at supporting the use of pen and paper. Instead, the tablet is used directly interact with the digital system, e.g. for drawing in a graphics application or for positioning the mouse pointer. However, the induction principle still works if one or several sheets of paper are positioned between the graphics tablet and the stylus. For this reason, they can be used to track pen input on real paper. This approach enables very high resolutions (about 1000 to 5000 dpi). However, it has the same drawbacks as ultrasonic tracking. The user must manually calibrate the paper sheet and must indicate page changes. Moreover, the interaction is restricted to the small surface of the tablet. Tracking interaction on paper with underlying graphics tablets is a common approach for research prototypes (e.g. [MPL<sup>+</sup>02, EMBW08, Tay08]).

**Absolute position tracking.** A second class of approaches tracks the absolute position of the pen, i.e. its position with respect to the paper sheet and not with respect to an external device. This requires that positional information is directly encoded on the paper sheets. It is the digital pen that decodes the position, which makes external devices and calibration obsolete. As the position which is encoded on the physical sheets can also contain a page identifier, the pen is able to detect on which page it is used. Hence, it is not necessary to manually indicate page changes or on which page one is currently using the pen. Users can therefore work very naturally with multiple sheets of paper.

*Anoto digital pen and paper* [Anob] is the currently most widespread solution of this class. The positional information is encoded on the paper sheets with a two-dimensional dot pattern (Fig. 3.6). This is printed on the page using an ordinary laser printer. Each dot has a position on an imaginary grid which overlays the page.



Figure 3.7: Anoto pens. Left: Nokia SU-1B (© Nokia). Right: Livescribe Pulse Smart-Pen (© Livescribe Ltd.)

By slightly displacing each dot in one of the four directions, a single dot encodes 2 bit of data. Each 6x6 matrix of dots encodes a unique position in the Anoto coordinate space. This indicates both the identifier of the page and the position on that page. The pattern is slightly visible to the human eye, making the print product appear slightly gray like recycled paper.

The Anoto pen behaves like an ordinary ballpoint pen on paper. If the print product contains the pattern, it is read and decoded by the pen, which therefore includes an infrared camera and a processing unit. The Anoto technology allows a resolution of about 850 dpi. In addition to the position in the two-dimensional coordinate space, current pens register pen tip force and timestamps. The technology further allows to detect the rotary and tilt angles of the pen (although this is not supported by the firmware of current pens). An Anoto pen is depicted in Fig. 3.7 (left). Depending on the capabilities of the pen, data is temporally buffered on the pen until it is synchronized with a computer via USB or it is streamed in real-time to a nearby computing device using a Bluetooth connection. Anoto pens include a battery and can therefore be used in mobile conditions.

*Anoto pens* are produced by several manufacturers. Current models are Logitech io2 (now rebranded Anoto io2), Nokia SU-27W and Maxell DP-201 (now rebranded Anoto DP-201). These standard models do only to capture and digitize pen traces to transfer them to a computer. It is the computer that further interprets the strokes and it is not possible to provide system feedback to the user directly on the pen. Liao et al. [LGL06] extended one of these pen models with auditory, tactile and visual feedback and explored with that prototype several pen-based feedback mechanisms.

Recently, several new pen models have been commercialized that include addi-

tional processing capabilities and output devices. These pens are able to interpret the strokes and react to user input. The Fly Fusion Pentop Computer [Fly] has a built-in speaker. It reads out the translations of handwritten words, can be used as a calculator and plays MP3 files. The Livescribe Pulse Smartpen [Liv] (Fig. 3.7 right) has a built-in microphone, speaker and an OLED display. In addition to capturing notes, it can be used to record and playback audio and as a calculator. An SDK provides for developing further applications (so-called penlets) which are executed directly on the pen. While these novel pens enable to develop more interactive interfaces including direct system feedback on the pen, they do not provide a wireless Bluetooth connection. For this purpose, interactive applications are currently restricted to a local use on the pen.

Recent research demonstrated how Anoto pens can also be used for pen *input on rear-projection screens* [BHH<sup>+</sup>07]. This approach relies on printing the Anoto pattern on a translucent foil. The digital content is projected on this foil using a commercial beamer. As on paper, the pen decodes its position from the printed pattern. This position can be translated to the position on the screen. This enables to produce very large pen-enabled displays that have a high input resolution at relatively low cost.

Another approach encodes positional information on paper using a grid of linear barcodes that are printed on the document with conductive ink [Sig06, p. 26]. This grid is slightly visible to the human eye. An inductive pen measures the inductivity and decodes it to a two-dimensional position. In contrast to the Anoto technology, which requires expensive optical components, inductive pens can be produced at a much lower cost. However, the resolution is not high enough to capture handwriting but only allows to distinguish larger hot-spot areas on a document.

### Electronic Components on Paper

The technologies for capturing user input, which we have discussed above, all keep paper as a purely physical medium (possibly expect for attaching a passive RFID tag). The electronics for digitizing or tracking is contained in specific devices which are not part of the paper sheets. This has the advantage that paper can be printed as usually.

Other approaches augment paper by attaching or embedding electronic components. This enables a greater diversity of interactions, but these augmented paper documents are more complicated and more expensive to produce.

An example is *Pulp-based computing* [CHBM07]. The authors embed diverse components directly into paper during the papermaking process. These components comprise microphones, bend sensors, LEDs, speakers and vibrating motors. *Voodoo-Sketch* [BHG<sup>+</sup>07] is an ad-hoc physical interface toolkit. This allows to plug various electronic controls, such as push buttons, switches and sliders, onto paper palettes, which have embedded conductive layers. Moreover, it is also possible to sketch controls using an Anoto pen. The *Computational Sketchbook* [BHE09] follows a similar approach. The user can create interactive paintings by attaching electronic compo-

nents onto ordinary paper. These components comprise speakers, motors, switches, LEDs and batteries. The user establishes the conductive connections by painting them with conductive ink. The electric circuit thus becomes a visible part of the artwork.

Our research does not follow this approach, as our goal is that it remains easy and cheap for the end-users to print their documents.

### 3.3 Pen-and-Paper Applications

Having discussed basic technologies for digitizing the contents of paper documents and the interaction with them, we now review applications which include a Pen-and-Paper User Interface and the interaction techniques used in these applications.

We distinguish two main classes of applications. A first class focuses on writing on documents. Digital pens are well-suited for this purpose because it is very easy and intuitive to write with a pen on paper. These applications enable the user for example to make handwritten notes or annotations on paper that are automatically digitized (e.g. [MPL<sup>+</sup>02, Gui03, YLK<sup>+</sup>06, Wel93]), to structure documents with handwritten links and tags (e.g. [HHT99, AAH97, LGHH08]) or to fill in printed forms [Anoa, VV06]. A second class of application leverages the mobility and flexibility of pen and paper for purposes other than writing. This includes using pen and paper as an interactive tourist guide [Sig06, p. 153 sqq.] or as a remote control for TV sets [HKP08].

In the remainder of this section, we will present and discuss applications and interaction techniques of the first class, as the focus of this work is on writing on documents. Following the three main activities which are supported by CoScribe, we distinguish applications and interaction techniques for notetaking and annotation (Sections 3.3.1 and 3.3.2), for cross-media hyperlinking (3.3.3) and for paper-based tagging (3.3.4). We present the functionality provided by prior applications, the underlying interaction metaphors and principles, how information on paper and on digital media are perceptually coupled and how collaboration is supported. First, the applications are described in a neutral manner – providing a general state of the art survey –, before we give a critical discussion with our requirements in mind in Section 3.4.

#### 3.3.1 Paper-based Notetaking

The main principle of augmented paper notebooks is to allow users to make handwritten free-form notes and sketches on a notebook of real paper sheets. These are automatically digitized and are available in a software viewer. This basic functionality is typically included in software for end-users that is shipped with the digital pen, e.g. with the *Pegasus Tablet Notetaker* [Peg] (Fig. 3.5 left) or the *Logitech io2 Anoto pen* [Log]. In contrast to other means for enabling pen input, such as graphics tablets or Tablet PCs, a physical notebook has the advantages of paper, e.g. for browsing through the physical notebook using both hands.

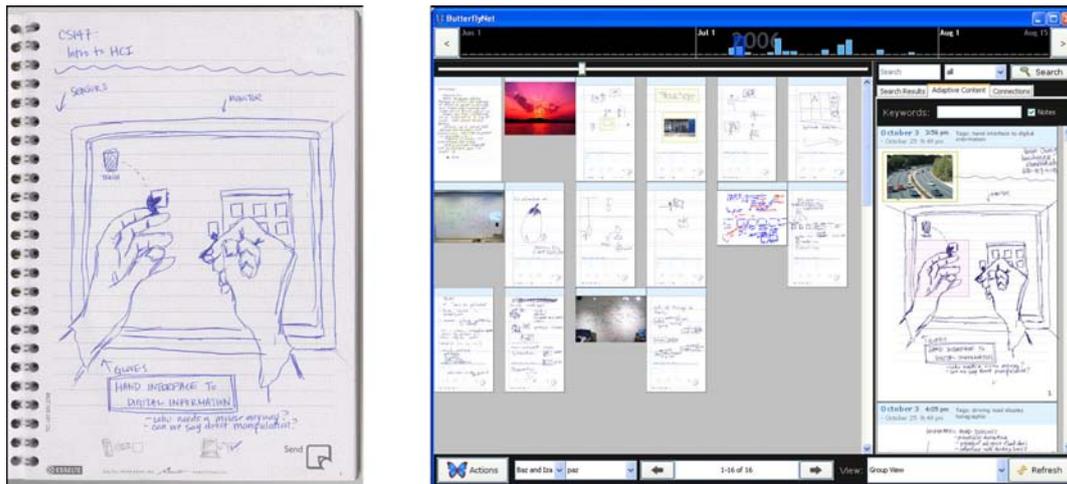


Figure 3.8: The *iDeas* notebook [LMK<sup>+</sup>06]. Left: Paper notebook. Right: Shared digital notebook including selected pages of multiple users

**Additional media.** More advanced research prototypes demonstrate how these basic paper and digital notebooks can be augmented by additional media. Yeh et al.'s *ButterflyNet* [YLK<sup>+</sup>06] is an augmented notebooks for biologists. A smart camera enables the user to take photos in the field. These can be embedded in the digital version of the notebook by selecting an appropriate position in the paper notebook with a pen gesture. *Memento* [WQK07] is a digital-physical scrapbook that uses a similar concept for embedding pictures or videos. The *Audio Notebook* [SAS01] closely couples handwritten notes on a paper notebook made during a meeting with an audio recording of that meeting. By tapping with the pen on a note, the recording is played back at the point in time when that specific note was made. This is a very intuitive and powerful method for spatially indexing temporal data, which is also included in the commercial Livescribe pen [Liv]. These systems have inspired our approach of linking paper documents and digital media.

**Collaborative applications.** Several augmented notebooks leverage the capabilities of computing for supporting group work. The following concepts can be distinguished:

A digital version of one's own notebook can be viewed by other persons over a network connection. In the most simple case, the user cannot select passages to share but the entire notebook is made available to other users. This is the case with *Memento* which allows to access shared notebooks in a standard Web browser. The *iDeas* notebook [LMK<sup>+</sup>06] (Fig. 3.8), an enhanced version of *ButterflyNet*, offers similar functionality. Moreover, users can add selected contents of their personal notebooks to a shared group notebook.

A second concept reported in the literature supports co-located collaboration.

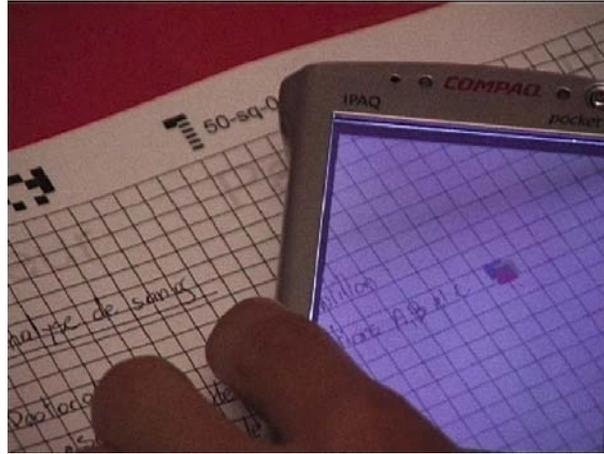


Figure 3.9: Interaction Lens of the A-Book for accessing digital content on the physical notebook [MPL<sup>+</sup>02]

*AirTransNote* [MKS07] focuses on communication within the classroom. Students take handwritten notes on a paper notebook. The notes are digitally captured and transmitted to the PC of the instructor. This enables the instructor to provide feedback and to discuss notes of students by projecting them on a screen. Moreover, the system uses handwriting recognition and clusters the recognized notes of all students. This enables the instructor to quickly get an overview on the notes.

**Closer perceptual coupling of printed and digital representations.** In all systems discussed above, there exist a large gap between printed contents on paper and digital contents that can only be accessed on a separate computer. This spatial and conceptual distance turns attention from the notebook to the computer, which moreover requires different interactions (e.g. keyboard and mouse instead of the pen). Therefore, physical input and digital output are perceptually distinct. We will now discuss approaches that provide a closer coupling of both media.

The *A-Book* [MPL<sup>+</sup>02] uses a graphics tablet which is positioned below the paper notebook to capture notes and sketches. Users can access the digital version of the notebook not only in a separate software viewer but also directly on the paper notebook. For this purpose, it leverages the metaphor of an interactive lens. Similar to a physical lens, the user can place a PDA on the notebook (Fig. 3.9). The position and orientation of the PDA is automatically tracked. This enables the system to display on the screen of the PDA a digital version of the area of the underlying paper notebook which is covered by the PDA. The illusion of a transparent PDA is maintained for any orientation of the PDA. The PDA overlays the contents of the paper document with additional digital information, such as hyperlinks. This seamlessly integrates physical and digital information, but only on the small screen of the PDA which must be moved over the page to successively view the digital

contents of the entire page.

Augmented desks and augmented walls support larger surfaces. As described above, the contents of paper documents are tracked using a camera and a beamer projects digital information onto the same surface. This principle is for example used in the *Digital Desk* [Wel93]. An example application enables users to make a sketch on paper and to copy this sketch to a digital collage. The digital contents are visible on the same surface. Moreover, the user interacts with both printed and digital contents using the same stylus.

A further approach is providing access to a digital version of notes via text-to-speech on an audio channel. *EdFest* [Sig06, p. 153 sqq.] is a mobile interactive guide containing a printed guide, an Anoto pen and a headset. The guide includes empty pages for writing handwritten reviews of individual events. These notes are transmitted to a central sever. Another user requesting information on that specific event can then play back this review on the headset.

### 3.3.2 Paper-based Annotation

This section discusses applications and concepts that allow the user to annotate existing documents instead of taking notes on empty pages. A large number of systems support annotations on digital documents. Typically, these are typewritten annotations, for example in word processors (e.g. Microsoft Word), document viewers (e.g. Adobe Acrobat) and web browsers (e.g. Annotea [KK01], Diigo [Dii]). Some systems use pen input for handwritten annotations of digital documents (e.g. Microsoft Journal for Tablet PCs) and include functions for sharing notes with collaborators (e.g. Notepals [DLC<sup>+</sup>99] and Livenotes [KWI<sup>+</sup>05]). However, these approaches cannot fully imitate the advantages of paper. For example, it is easier to get an overall sense of the structure of a document by laying out its pages on the desk than by sequentially browsing through its pages on a single screen. A considerable number of research prototypes have been published that support handwritten annotations on printed documents.

**Annotating printed versions of digital documents.** An important system of this class is *Paper Augmented Digital Documents (PADD)* [Gui03]. This system introduced the main principle of paper-based annotation. A digital document is printed on paper. This printout is used as a proxy to interact with the underlying digital version of this document by making handwritten annotations on paper. These annotations are automatically digitized and added to the digital document. The digital document is accessed on a computer. Moreover, the user can print updated versions of the document including new annotations. This results in a digital-paper-digital annotation lifecycle. PADD does not interpret the handwritings but simply captures and visualizes them. *PaperProof* [WISN08] demonstrates how handwritten correction marks can be automatically interpreted to modify the underlying digital document. Using handwriting recognition and a module that parses the symbolic contents of

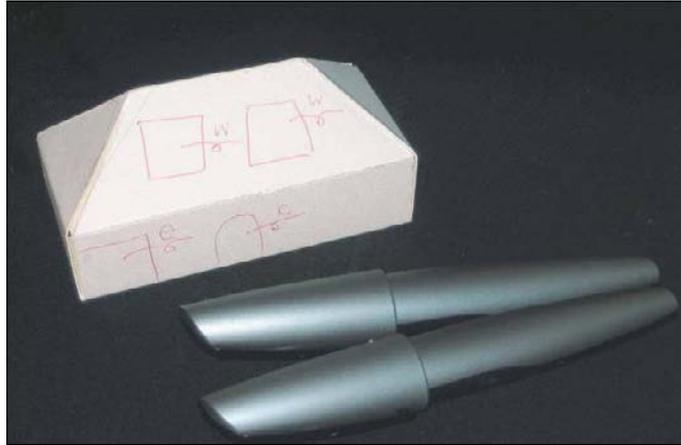


Figure 3.10: ModelCraft provides for annotating three-dimensional paper models [SGHL06]

the document, it is for example possible to delete text by striking it out or to insert text by writing it at the appropriate position.

PADD and PaperProof focus on standard documents in knowledge work. However, the same principle applies to more nonstandard documents, such as large posters [YBK<sup>+</sup>06], maps [MHBC08] or musical partitions [TLM09]. Moreover, related research [SGHL06, SGAL07] demonstrated how Anoto pen and paper can be leveraged to annotate three-dimensional physical models made of paper (Fig. 3.10). The annotations are automatically added to the underlying CAD model. This shows the versatility of paper annotations. However, in contrast to flat documents which can be easily re-printed, creating updated versions of physical models requires manual work. This limits the paper-digital annotation lifecycle.

**Hybrid printed and digital collages.** Another approach of paper-based annotation does not use a printed version of a digital document as a proxy for adding annotations to it. Instead, the digital document is displayed on a screen. Handwritten annotations are added by attaching a paper sticker, which can be annotated, onto the screen. This creates a hybrid physical and digital collage with all contents being available in the digital version of the document. The *Designer's Outpost* [KNF<sup>+</sup>01] focuses on creative planning tasks with Post-it stickers (Fig. 3.11 left). It is a rear-projection smartboard onto which users can make digital handwritings with a digital stylus. This digital document can be annotated by attaching physical post-it stickers containing arbitrary handwritings and drawings. The positions and the contents of the stickers are automatically digitized and added to the digital document. *DigiPost* [JYWS07] uses a similar principle on a horizontal tabletop display.



Figure 3.11: Left: The Designer's Outpost [KNF<sup>+</sup>01]. Right: PaperPoint [SN07]

**Collaborative applications.** Neither of these systems does offer specific functionality for group work. The following two systems show how annotations on paper can support co-located collaboration. *PaperCP* [LGA<sup>+</sup>07] supports in-classroom collaboration in a way similar to AirTransNote. Students can annotate printed handouts of lecture slides and electronically transmit their annotations to the instructor. The paper discusses how students can define own notes to be excluded from sharing. It proposes declaring different regions on a paper page to be exclusively used for private or for public notes, using different pens for private or public notes and differentiating with pen gestures. *PaperPoint* [SN07] is a paper-based interface for controlling slide presentations from a printout and annotating slides on paper (Fig. 3.11 right). All annotations made on the slide printout are added to the digital slides in real-time. Several printed buttons allow to switch between slides during a presentation.

Recent research [SGF<sup>+</sup>09] examined how users can interact with a digital pen that has a built-in projector. The pen projects a superimposed digital information layer onto the paper document. This provides for selecting items in digital pie menus and for accessing annotations created by other users. However, this research is a pure conceptual study. As current projectors are still too large to be applied to a pen, the authors simulated a projecting pen. For this purpose, they tracked the location and orientation of the pen using visual camera-based tracking and projected the superimposed layer with a standard projector that was mounted at the ceiling above the table.

### 3.3.3 Paper-based Hyperlinking

#### Token-based Access to Digital Resources

One class of physical hypermedia offers a physical way for accessing digital documents and applications. Each digital resource to access is represented by a physical



Figure 3.12: A WebSticker (small photo) contains a barcode that is associated with a Web resource [LRH00].

object. By manipulating this object (e.g. by holding it near to a barcode reader or to an RFID reader or by pressing a push button on that object), the associated digital resource is accessed. This provides an easy and intuitive way for selecting and opening documents or applications. The physical proxies can be managed and ordered by spatially arranging them in the physical world. Moreover, they can be easily shared with co-workers. An example is *WebStickers* [LRH00] which uses paper tokens that contain a barcode (Fig. 3.12). The user can associate a token with a single Web page or with a collection of Web pages. By holding the token under a barcode reader, the associated resource is displayed on a computer screen. A similar approach has recently been commercialized (e.g. [Bee, Sem]). Two-dimensional barcodes (e.g. on advertisements, stamps and the like) act as links to Web pages. The barcode is decoded using a standard mobile phone that has a built-in camera. Based on a similar approach *Palette* [NInPA99] follows a similar approach and uses barcode-enhanced paper cards for accessing individual slides of slide presentations. *PaperButtons* [nPSN00] overcomes the limitations of barcodes that must be placed below a barcode reader or in front of a camera phone. It extends the *Palette* system by replacing the barcode on a paper card with a push button. A unique ID is transmitted to the system over a wireless connection when this button is pressed. *Augmented Filing Systems* (e.g. [SJH<sup>+</sup>08, JM09]) track the physical location of folders in a filing cabinet and synchronize these with their digital counterparts. *DigiClip* [DBEK04] is an electronically augmented paper clip which includes LEDs, sensors and wireless communication. The user can attach a clip to a paper document. Since the system tracks the position of the paper clips, it is easy to find the physical copy of a digital document later on.

This linking between a physical and a digital instance of the *same* resource is not the focus of this thesis. Although in *CoScribe*, printed representations of a docu-

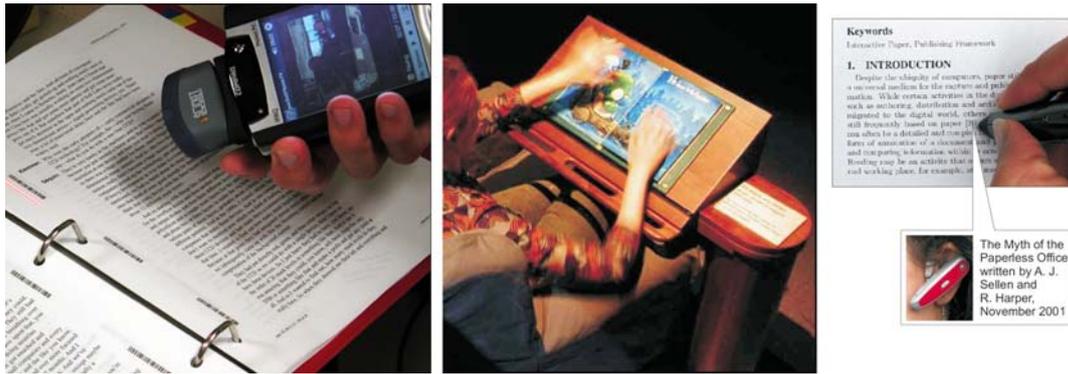


Figure 3.13: Three systems supporting pre-programmed hyperlinks. Left: Books with Voices [KGWL03]. Center: Listen Reader [BCG<sup>+</sup>01]. Right: Print-n-Link [NSW06b].

ment can also be used for accessing the corresponding digital representation, the main function of printouts is not being a simple handle but being a complete and self-contained representation of the document. We understand hyperlinks as references between *different* resources. In the following, we will review approaches for hyperlinks between printed documents and digital media.

### Preprogrammed Hyperlinks between Documents

A number of systems enable the user to access digital resources by selecting hyperlinks on a printed document. The aim of these systems is to support users in reading documents and accessing additional digital information if necessary. We can classify these systems following the technology which is used to select hyperlinks on paper (and the resulting interactions), the type of the printed document as well as the type of the digitally linked media and the device on which digital media is made available. The underlying interaction metaphor of all these systems is selecting a link hot-spot with a deictic gesture for accessing the associated digital resource. Depending on the technology, the deictic gesture is tapping with a stylus or the finger, clicking with the mouse or scanning with a barcode reader.

*Books with Voices* [KGWL03] illustrates well the principle of interactive books. A book which is printed on ordinary paper is augmented by cross-media hyperlinks to digital videos. For this purpose, a linear barcode is printed in the margin of the document at each place where a hyperlink starts. When the user wants to follow this hyperlink, she scans the barcode with a PDA that is augmented by a barcode reader (Fig. 3.13 left). The video is then played back on the PDA.

While this relatively simple technology is easy to deploy, it requires a considerable number of printed barcodes, which can interfere with the visual appearance of the printed artwork. Moreover, it is not possible to define a link hot-spot directly

within the document but only in its margin. For these reasons, the following systems rely on printed barcodes only for identifying the current page but use other means for identifying the position on a page which is selected by the user.

*ActiveBook* [SMF99] makes areas of a printed book “active” in the sense that the user can select an area to access an associated Web resource on a nearby computer. An area is selected by positioning a device similar to a mouse on it. This device tracks the user’s pointing position relative to a reference location within the paper document. As this technology does not detect link hot-spots using visual information on the document but with the mouse device, *ActiveBook* does not require marking up link hot-spots on the document with barcodes or the like. Moreover, link hot-spots can be of various shapes and sizes, allowing very flexible link anchors.

*Animated Paper Documents* [RSW<sup>+</sup>97], an extension of the Digital Desk [Wel93] enables the user to print Web pages with their hyperlinks. The user can select a hyperlink on a printed document by tapping on it with a stylus, whose position is tracked by a camera mounted above the desk. The linked document is then displayed in a browser window which is projected onto the desk. This creates a very close coupling of printed input and digital output but requires a complex technical setup that restricts the mobility which is inherent to paper.

While both *ActiveBook* and *Animated Paper Documents* require that barcodes are printed on each page in order to identify this page, the following systems do not use barcodes at all. This leaves full freedom for the design of the printed artwork. The *Listen Reader* [BCG<sup>+</sup>01] (Fig. 3.13 center) is an augmented physical book with embedded hyperlinks to audio files. The page of the book which is currently open is detected using RFID technology. A passive RFID tag is embedded in each page of the book. An RFID reader in the back cover of the book detects which pages form the right hand side of the opened book. Pages on the left hand side are outside the range of the reader. This allows to infer the page which is currently open. Furthermore, *Listen Reader* does not only detect touch events to select link hot-spots, but also the distance of the user’s hands to the book. By altering this distance, the user can control the volume of the audio playback. For this purpose, the *Listen Reader* relies on electromagnetic sensing. Four electrodes, one in each quadrant of a page, create four electromagnetic fields. A sensor detects touch and hovering interactions that the user makes with her hands.

*Print-n-Link* [NSW06b] aims at reducing the burden of retrieving articles that are referenced in scientific publications. When the user prints a scientific article, *Print-and-Link* automatically detects the references within this article. The text of each reference (e.g. “[12]” or “(Author, 2007)”) becomes a link hot-spot for the referenced article. By tapping twice with the Anoto pen on a link hot-spot on the printed document, the referenced article can be electronically accessed (Fig. 3.13 right). Depending on the user configuration, a nearby computer displays the PDF of the article or shows its metadata.

These systems offer appropriate support as long as the focus is on reading a document and it is sufficient to have access to some pre-programmed digital re-

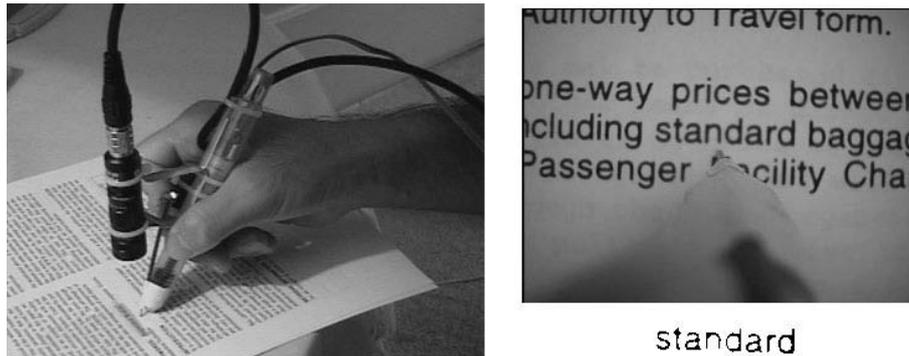


Figure 3.14: PaperLink uses a pen that is augmented with a camera (left). From the camera image, a text pattern is extracted (right) that can act as an anchor for a hyperlink [AAH97].

sources. If the user has a more active role and wants to create own hyperlinks, pre-programmed links are too restricted.

### User-defined Hyperlinks

Another class of systems allows users not only to follow (pre-defined) hyperlinks, but also to create and follow personal hyperlinks. This was already the aim of early hypertext visions. Vannevar Bush's Memex [Bus45] was the vision of an electro-mechanical machine which stores all books and other documents of a user on microfilm. Memex enables the user to read and annotate these documents as well as to create associations (we would call them hyperlinks today) between these documents: "The process of tying two items together is the important thing." (Vannevar Bush, *As We May Think* [Bus45]). These associations provide a systematized access to information that is distributed between several documents. Bush influenced Ted Nelson's concept of hypertext [Nel65]. In contrast to the simplified hypertext model of the World Wide Web and similar to Bush, Nelson envisaged that users can create their own hyperlinks and annotations. (such as those by Vannevar Bush [Bus45] and Ted Nelson [Nel65]). However, these approaches constrain the user to work with documents that are displayed on a machine and not on paper. We will now review systems that support hyperlinks on *physical* documents. Each of them uses a different technology for bridging physical paper to computers, demonstrating the variety of approaches.

*PaperLink* [AAH97] is based on a pen with a built-in camera that scans an image of the area of the document around the pen tip (Fig. 3.14). A specific text pattern (i.e. a mark on the document scanned by the pen) can be associated with a digital resource. When this mark is returned to later on and scanned with the pen, the associated resource is displayed on a nearby computer. The *Interactive Multimedia Textbook* [LCC07] offers similar pen-based interactions for creating and following

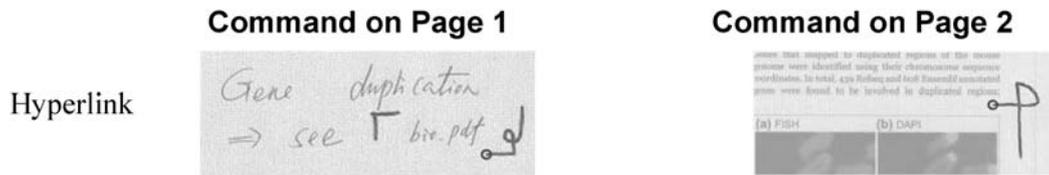


Figure 3.15: Pen gestures for creating a hyperlink in PapierCraft [LGHH08].

hyperlinks from printed documents to Web pages. Since it uses an ultrasonic pen (see Section 3.2.3), link anchors are not bound to specific marks on the document (e.g. an individual word) but to freely-defined hot-spot areas. Tapping with the pen on such an area displays the associated digital document. Both of these systems are limited to links from paper to digital media.

In contrast, *PapierCraft* [LGHH08] enables the user to create hyperlinks between two paper documents. It is a pen-based command system for paper documents, relying on the Anoto technology and containing a rich set of pen gestures. These gestures provide for creating and following hyperlinks on documents, for copy/paste within paper documents, for tagging documents, for performing Google searches using keywords from the printed document and for mailing portions of a document to other persons. A hyperlink is created by drawing two specific pen gestures on both link anchors (see Fig. 3.15). For defining the extent of the link anchors, *PapierCraft* provides flexible ways, e.g. underlining or encircling text, selecting several lines of text with a margin bar or specifying a rectangular area. For links from or to digital documents, the same gestures can be used on a Tablet PC on which the digital document is displayed. Later on, a hyperlink can be followed in the document viewer on a computer, but not from paper. In contrast to our work, *PapierCraft* is limited to PDF documents and does not support links from and to Web pages.

Unlike the systems presented above, which model hyperlinks as being part of the physical layer of the printed document, the *A-Book* [MPL<sup>+</sup>02] treats hyperlinks on paper content as being part of a superimposed digital layer. The *A-Book* is an augmented notebook for field biologists and has been described in Section 3.3.1. In order to create a hyperlink on the paper notebook, the user would place the Interaction Lens onto the notebook at the location where the first link anchor should be. Then, she would tap with the pen on the display of the Interaction Lens at that precise location. When the pen is kept held down for a short moment, a link anchor is created and visualized in the Interaction Lens. The second link anchor would be created on the same or on another page using the same technique. A hyperlink can be followed by placing the Interaction Lens at the appropriate position onto the notebook and tapping with the pen on a hyperlink symbol which is displayed as part of the superimposed digital layer. An advantage of the *A-Book*'s linking technique is that the Interaction Lens keeps a list of open hyperlinks. Therefore, it is not necessary to immediately define the second link anchor.



Figure 3.16: The PaperPDA [HHT99]. Center: A form for writing e-mails on paper. Right: Physical stickers for creating hyperlinks.

The *PaperPDA* [HHT99] incorporates an entirely different interaction technique for creating links than the deictic tapping or clicking gestures of the systems discussed above. It combines a conventional paper notebook, calendar and organizer with electronic support. Using specific pre-printed paper forms, the user can for example take notes and write e-mails using real pen and paper (Fig. 3.16). This is particularly helpful in mobile settings. Back in her office, the user scans all new or updated forms using a desktop scanner. Specific visual marks on each form allow the system to identify the orientation and the type of the form. PaperPDA then detects marks that the user has made on interface elements (such as printed check boxes) and uses OCR for recognizing text. The PaperPDA system then automatically performs the electronic operations requested by the user, e.g. sending an e-mail. Hyperlinks between two paper documents are created with physical stickers (Fig. 3.16 right). The user attaches two corresponding stickers to the locations of both link anchor. On each of these stickers, an ID is printed that enables the system to detect corresponding stickers when the forms are scanned. This automatically creates a digital version of the hyperlink. These physical stickers were a source of inspiration for our tangible bookmarks.

*DocuDesk* [EMBW08] is an augmented desk system that supports the user in the combined work with printed and digital documents, focusing particularly on quickly re-establishing the state of open windows and documents when resuming a task. It provides a very intuitive way for creating many-to-many links between printed and digital documents that are positioned on a pen-sensitive graphics tablet. A link is established not by defining two separate anchor points but by connecting two documents by drawing a line. This interaction emphasizes on the connecting character of hyperlinks. *DocuDesk* relies on a 22 inch pen-sensitive Wacom screen in a horizontal tabletop configuration. Digital documents are displayed on that screen. Moreover, the user can place physical documents on the screen (Fig. 3.17). Their positions are tracked by a camera which is mounted above the desk. The camera also captures images of the contents of documents and of the annotations that are made with a traditional pen. Hence, the system has full knowledge not only of open

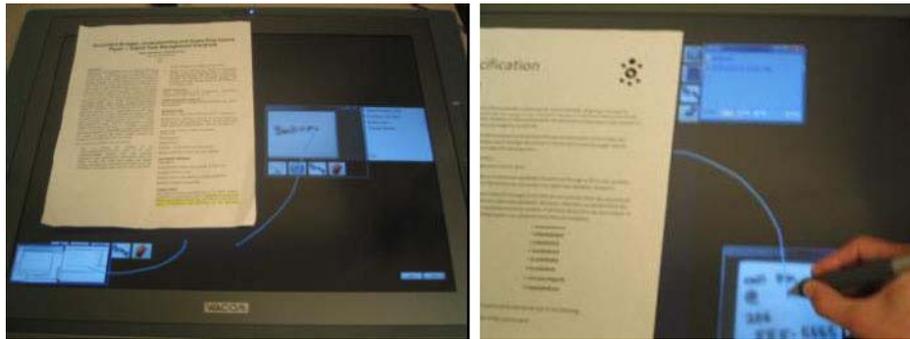


Figure 3.17: The DocuDesk [EMBW08]. A hyperlink is created by drawing a line on the pen-sensitive display that connects the documents (right).

digital applications but also of the contents and the locations of paper documents. This enables quickly re-establishing this state later on (with having digital versions of the printed documents displayed on the screen). A many-to-many link is created by drawing connecting lines on the pen-sensitive tabletop display between all documents to be part of this link, regardless if the document is a printed or a digital one (Fig. 3.17 right). However, links apply only to entire documents. To follow a link later on, it suffices to place one of these documents onto the screen. All other documents are then displayed on the screen, surrounding this document and being connected by lines.

While these interaction techniques for creating and following hyperlinks are varied – deictic pen gestures, connecting lines, physical sticking actions – , they all have an important property in common: The user performs an action on or with the documents, but the location of the documents keeps unchanged. We now discuss a fourth technique which relies on collocating physical objects in order to establish a hyperlink between them. A good example for this technique is the work by Grøn-bæk et al. [nKrE03]. This introduced the term *Physical Hypermedia* for hyperstructures that contain both physical and digital objects. A prototype supports users in linking and grouping RFID-tagged physical objects by holding them near an RFID reader. A software browser visualizes the hyperstructure in a 2D view by placing associated objects near each other. However, as the focus is not on paper documents but on general physical artifacts, a more specific document support, (e.g. linking document passages instead of entire documents) is missing.

### 3.3.4 Paper-based Tagging

On the Web, tagging resources is nowadays very common. Personal tagging is supported by browser bookmarks. Collaborative bookmarking tools such as del.icio.us<sup>1</sup>,

---

<sup>1</sup><http://del.icio.us>

digg<sup>2</sup> or cite-u-like<sup>3</sup> enable sharing own bookmarks with other users. Many Web 2.0 platforms contain their own tagging functionality, e.g. for photos (flickr<sup>4</sup>) or videos (YouTube<sup>5</sup>). However, these approaches are restricted to resources that are available on the Web. Moreover tags apply always to entire documents and it is not possible to tag sub-passages of them.

Before we review in more detail systems for tagging paper documents, we will have a look at digital notetaking and annotation systems that inspired our work. Some collaborative forum and annotation systems (e.g. [LTZ05]) enable the user to categorize his or her contributions following their functions in the discourse. This relies on the concept of speech acts [Aus62, Sea69, Win86] and is called “scripting” [Kin07]. This classification of individual annotations or individual forum posts inspired our approach for tagging individual annotations with semantic categories.

We now review approaches for tagging contents on paper documents. Only little work has explored this issue. Some of the *commercial software packages* that are shipped with Anoto pens (e.g. Logitech io2 software [Log], Oxford Easybook M3 Digital Notepad [Oxf]) allow the user to mark up the notes in order to control further electronic processing steps already while taking notes on paper. This comprises typical office tasks, such as adding a new task to Microsoft Outlook, inserting a note into a new document in Microsoft Word or sending a note as an e-mail to another person. In order to invoke the electronic command, the user writes a specific character or symbol in the margin of the notebook and selects the operand by drawing a vertical line beneath the appropriate note(s). However, this activity does not create tags in a sense that the notes are categorized for later retrieval.

Semantic tagging in a more precise sense is supported by *PapierCraft* [LGHH08], which has been already introduced in the previous section. It is a gesture-based command system that offers two methods for tagging contents on printed PDF documents. On the one hand, the user can tag contents using predefined categories. This is done with a specific pen gesture for predefined categories (Fig. 3.18 upper row). The cardinal direction of the ending of the pigtail gesture decides upon the category. The current prototype supports two predefined categories. This approach guarantees readability on paper and a reliable gesture recognition. However, only a small number of tags can be supported (up to the eight cardinal and secondary directions). Moreover, the gestures are abstract, since they have no intuitive connection to the category. The user must memorize the gestures and the corresponding keywords. On the other hand, in addition to predefined categories, the user can tag contents with freely-chosen keywords. In order to so, she performs a basic tagging gesture and adds a handwritten label (Fig. 3.18 lower row). In contrast to our work, the current prototype of *PapierCraft* requires an additional device, for instance a push button, for switching between a writing and a command mode.

In another publication, Liao et al. [LGA<sup>+</sup>07] discuss several ways for tagging in-

---

<sup>2</sup><http://digg.com>

<sup>3</sup><http://www.citeulike.org>

<sup>4</sup><http://www.flickr.com>

<sup>5</sup><http://www.youtube.com>

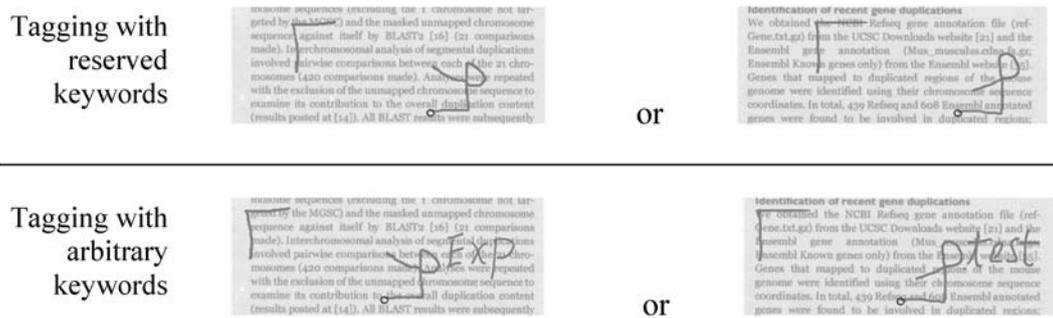


Figure 3.18: PapierCraft’s pen gestures for tagging document passages with predefined or freely-chosen keywords [LGHH08].

dividual notes with pen-based interactions. Spatial differentiation consists of providing separate areas on a document for different categories. The note is tagged with the category of the area it is written in. While this is a simple, quick and intuitive means, it forces the user to follow a predefined spatial arrangement for his or her notes and moreover supports only a rather small number of types. A second method, pen differentiation, means that the user switches between different pens whereby each pen is associated with a specific category. All notes written with that pen are tagged with this category. This does not interfere with the spatial layout of notes but switching between pens takes more time than spatial differentiation and moreover this requires additional hardware. In addition, research shows that students rather use one single pen than switching among many marking tools [Mar97]. A third solution consists of classifying notes by performing specific pen gestures, like the ones used in PapierCraft. This provides for a possibly large set of categories. However, current digital pens cannot recognize gestures by themselves. Therefore, gesture recognition is performed by the back-end system. In a mobile setting, where pen data is temporarily buffered on the pen, this has the disadvantage that the user cannot get real-time feedback on success or failure of gesture recognition. Novel generations of digital pens which include gesture recognition could solve this problem in the future.

Finally, *Quickies* [MM08] proposes using physical stickers for taking notes on paper, which are automatically digitized. Stickers can also be used to tag physical objects with a handwritten label. If the sticker is attached to the object, this can be retrieved in the physical space, as an RFID tag is applied to the reverse side of each sticker. While this approach is intuitive and powerful for tagging entire objects, it is not possible to tag sub-passages of documents.

| Requirement  | Supported by previous work? | Contribution of this thesis   |
|--|-----------------------------|---|
| R1 Support of both printed and digital documents                       | +                           | Richer variety of interaction metaphors   |
| R1a Similar interactions in the physical and in the digital worlds     | -                           | Unified pen-based interaction   |
| R1b Reliable interactions without digital feedback                     | +/-                         | Reliable interactions without the need for handwriting recognition  |
| R2 Supporting annotating, linking and tagging collections of documents | +                           | Richer variety of interaction techniques  |
| R3 Supporting co-located and remote collaboration                      | -                           | Concepts for paper-based sharing and integrated multi-user visualizations   |
| R4 Symbiotic integration into current information ecologies            | -                           | Integrated solution based on ecological perspective   |
| R4a Interactions that produce little extraneous cognitive load         | +/-                         | Easy and intuitive pen gestures from a small inventory  |
| R4b Flexible interactions  | -                           | Freedom to maintain current work practices; adaptable printed user interface; choice amongst interaction techniques |
| R4c Structured access to all information                               | -                           | Ecological view   |

Table 3.1: Overview of which requirements are supported by previous work

### 3.4 Discussion and Conclusions

In Section 3.1.3 above, we have already assessed prior work on *interaction models* of Pen-and-Paper User Interfaces. Existing models of the broader field of Tangible User Interface can be leveraged for describing the interplay between physical and digital representations and the relations between physical objects and the information they represent. However, we have identified a gap on the level of interactions, where existing models do not sufficiently describe Pen-and-Paper User Interfaces. We therefore contribute an interaction model of Pen-and-Paper User Interfaces, which will be presented in the next chapter.

This state of the art review has moreover shown that there exist a variety of *technologies* that provide for digitizing and tracking the contents of printed documents and the interactions with them.

In the remainder of this chapter, we will now provide a critical discussion of the contributions and shortcomings of the *applications* with respect to our design goals. Table 3.1 provides an overview of which of our requirements from Section 2.5 are met by this previous work and what are the contributions of this thesis to overcome limitations of previous work. The discussion is centered around the following four topics:

1. The interaction paradigms and metaphors used for the activities of notetaking/annotation, linking and tagging
2. The perceptual coupling of printed and digital documents
3. The support of collaboration
4. The extent to which the solutions can be integrated into existing information ecologies

### 3.4.1 Interaction Paradigms and Metaphors for the Conceptual Activities

Figure 3.19 gives an overview on the conceptual activities and on the types of collaboration that are supported by the applications. It shows that CoScribe, the contribution of this thesis, is the first application in this domain that covers remote asynchronous and co-located collaboration encompassing the conceptual activities of notetaking, annotation, linking and tagging.

**Notetaking and annotation.** Pure *notetaking* systems are not sufficient for our application domain, as the user should be able to annotate printed documents. We have shown that paper-based *annotation* systems are based on the two following interaction paradigms. First, a printout of a document is used as a physical proxy for interacting with a digital document. Second, physical paper stickers are used for annotating digital documents on a display. The resulting document is a hybrid physical and digital collage. The first class is an important source of inspiration for our work.

However, previous work on paper-based annotations forces the user to work with a predefined printed user interface. This stands in contrast to most applications with Graphical User Interfaces, whose interfaces can be adapted by the end-user. Our approach introduces Pen-and-Paper User Interfaces whose layout can be flexibly adapted by the end-user. Moreover, as will be discussed below, we contribute novel techniques for the asynchronous collaboration around paper-based annotations.

**Hyperlinking.** As far as the conceptual activity of *hyperlinking* is concerned, our goal is to let users create own hyperlinks between documents. Therefore, the approaches for token-based access and for pre-programmed hyperlinks are not appropriate. We have shown that the most widespread interaction metaphor for following hyperlinks consists of a pointing gesture that is performed on a link hot-spot. This is highly easy and intuitive and inspired our approach, which relies on deictic pen gestures for defining and following hyperlinks.

However, current approaches require the user to switch between different interaction devices for paper and for digital media. A further limitation of prior work is that links apply only to one fix scope. Link anchors are restricted either to a point at which the link starts or ends (PaperLink, PaperPDA, A-Book) or to an entire

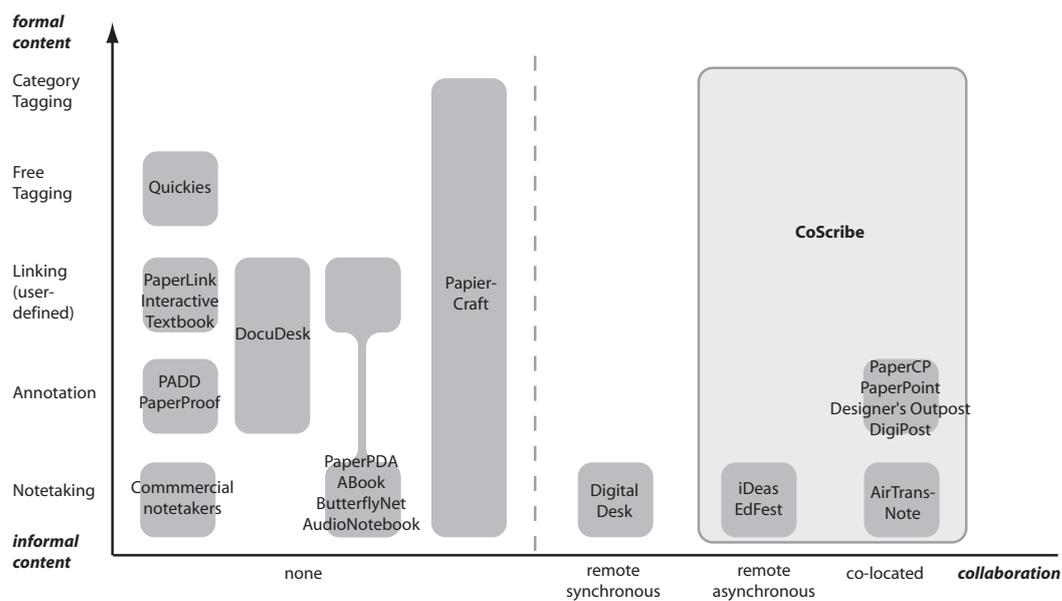


Figure 3.19: Overview of the applications in the design space: conceptual activities and collaboration

document without the possibility to reference sub-passages (DocuDesk, Physical Hypermedia). An exception is PapierCraft which provides for very flexible scopes. However, even in this system, scopes are restricted to one document and cannot span multiple documents. In contrast, CoScribe offers flexible link scopes from individual points up to collections of documents.

**Tagging.** PapierCraft is possibly the system most related to ours, as it provides for annotating, linking and *tagging* printed documents. However, PapierCraft does not offer links and tags on web pages and does not provide support for collaboration. Moreover, our approach has a different underlying interaction model. PapierCraft and the pen-based tagging approaches discussed in the literature rely only on interactions with *individual* pages. Users make gestures and write only on one single page. These interactions are not really specific to paper, as they would be possible in the same way on any pen-enabled display. Moreover, while this single-display mark-up approach is very flexible, it is only one facet of the rich interactions that are possible with paper. In contrast, CoScribe leverages such varied interactions as writing on pages, arranging and combining multiple sheets of pages and attaching stickers. Finally, the tagging functionality of PapierCraft relies on handwriting recognition. In mobile conditions, where current pens cannot provide real-time feedback on the success or failure of the recognition, this interaction is less reliable than our approach, which does not require handwriting recognition.

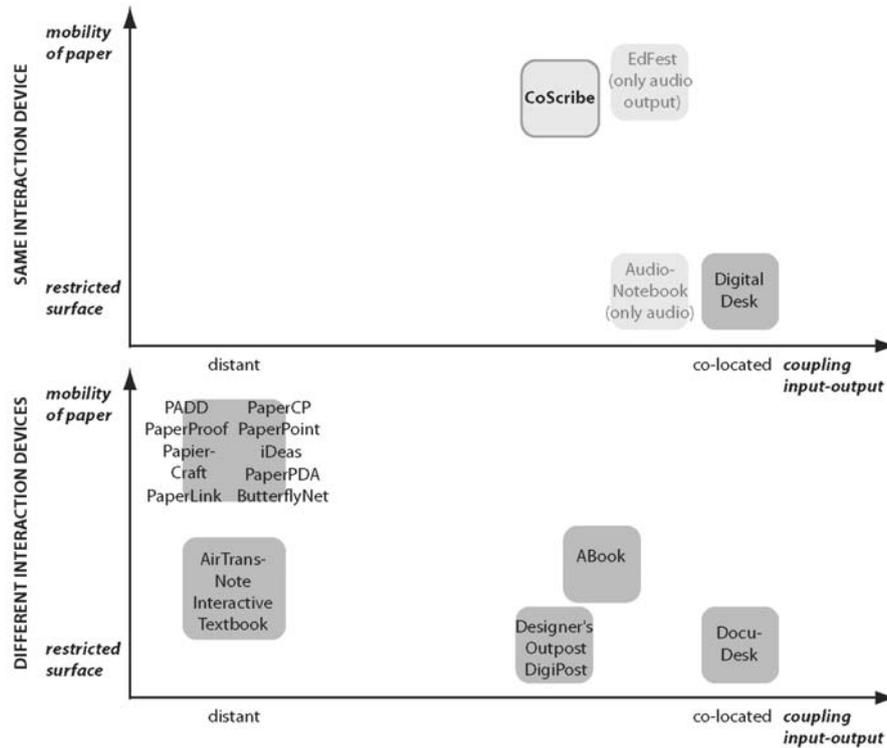


Figure 3.20: Perceptual coupling of paper and digital documents

### 3.4.2 Perceptual Coupling of Printed and Digital Information

Figure 3.20 systematizes related work according to the perceptual coupling of printed and digital information. It includes three dimensions: 1) the spatial proximity of paper-based input and digital output, 2) the extent to which the mobility of paper is maintained and 3) the aspect whether the same devices can be used in both worlds.

**Spatial proximity.** The x-axis illustrates the *spatial proximity of input on paper and digital output*, ranging from distant output on a nearby computer screen to co-located output realized by a projection of digital contents onto the paper documents. While audio output is perceived as near, it is a different modality which is not capable of providing visual feedback. Interactive lenses enable digital feedback which is spatially co-located with printed contents. However, they cover only small areas of a document. Augmented desk and wall systems display physical and digital contents on one single surface and provide support for large surfaces. This inspired our approach of using a tabletop display for digital information. In contrast, printed information is perceptually decoupled from digital information in most other applications. Digital information is only available on a distant computer screen and the user has to switch between different devices for interacting with both types of

media. This stands in contrast to a close integration of paper and digital media.

**Mobility of paper.** While augmented desk systems provide for a closer integration in this respect, they imply the disadvantage that camera tracking and projection restrict interaction to a small area below the camera and the projector. In contrast, our approach maintains the full *mobility of paper*. Only the interaction with digital contents is restricted to the tabletop display. This degree to which the mobility of paper is maintained is denoted by the y-axis.

**Unified interaction.** In order to further close the gap between printed and digital information, the user should be able to use the *same pen for all interactions* both on paper documents and on digital documents. No system except for the Digital Desk from related work fulfills this requirement. Most approaches clearly distinguish between the pen and paper modality and traditional interaction with keyboard and mouse when it comes to working with digital documents (Fig. 3.20 lower part). While PapierCraft allows to perform the same gestures as on paper on a Tablet PC, this requires switching from the Anoto pen to the stylus of the Tablet PC. Moreover, links can be followed only in the PapierCraft viewer and not by selecting a link hot-spot on paper. The A-Book also requires switching between a traditional pen for writing on paper and the stylus of the Interaction Lens. DocuDesk allows to use the same stylus for creating links on printed and on digital documents. However, one has to use a different pen for creating annotations on printed documents. While all functionalities of the Digital Desk can be performed with one single pen, the functionalities support only a small subset of our requirements. In contrast to related work, our approach provides for using the same pen for all interactions (annotating, creating and following links, tagging, sharing) both on paper and on digital documents.

Summing up, compared to most other digital pen and paper applications, which are characterized by a large gap between input on paper and digital output on a nearby computer screen and require different interaction devices, CoScribe provides a much more closer coupling of paper and digital media by leveraging a tabletop display and one single pen for both types of media. This creates one unified surface for both printed and digital documents.

### 3.4.3 Collaboration

A further dimension of this critical discussion is collaboration support. We have shown that paper-based systems mainly focus on individual users and do not provide support for collaboration. This is true for all systems that offer support for linking and tagging documents. As shown in Fig. 3.19, some notetaking and annotation systems include collaborative functionality. However, each of these systems covers only one collaborative setting. Moreover, the challenges of asynchronous sharing are not systematically addressed.

This includes privacy, which is an important issue of sharing notes with other users. The systems either ignore this issue (all notes are shared) or offer only limited support (select own notes to be added to a shared repository). Our work suggests a paper-based mechanism for sharing handwritten annotations, links and tags on various levels of visibility.

A second challenge is the visualization of shared content. We discuss how shared handwritten annotations of many users can be visualized in an integrated view. Moreover, we provide an ecological view that enables users to see at a glance which documents and pages of the collaborative document space are related and which links and tags are created by which user.

#### 3.4.4 Integration into Existing Information Ecologies

A final question is to which extent and how easy prior applications can be integrated into existing information ecologies. Prior applications support only a fraction of the *document types* that are used in typical information ecologies. These are either PDF documents, PowerPoint slides, Web pages or specific physical books. In contrast, our approach provides for linking not only PDF documents in their printed or in digital versions, but supports also PowerPoint slides, Web pages and all physical books which contain a EAN barcode.

Interaction techniques for annotating, linking and tagging should produce only *low extraneous cognitive load*. This is particularly important in situations which imply by their very nature a high degree of intrinsic cognitive load [CS91], such as university lectures. The interactions for annotation and notetaking fulfill this requirement if the underlying technology does not require the user to manually indicate page changes. The same is true for the concept of deictic gestures for following hyperlinks. However, the perceptual decoupling of printed and digital media, which we have identified above, leads to a higher degree of cognitive load than our approach, as users have to switch between different interaction devices. PapierCraft, the most significant previous work on paper-based tagging, requires to memorize rather complex pen gestures, which are not intuitive. These gestures may produce low cognitive load for power users. However, novices and users that utilize the system less frequently are very likely to be faced with high cognitive load. In contrast, our tagging techniques are more intuitive to use.

Moreover, the techniques should be *adaptable by the end-user* to make them fit the requirements of a given information ecology. Prior work does not allow the end-user to adapt the layout of the printed user interface. Moreover, the traditional practice of writing with a pen on paper is characterized by its extreme flexibility. The approach of this thesis maintains this flexibility. The user can continue making all informal comments, references, tags as before. Only if she desires that these contents are interpreted by the system, she performs a more formalized interaction, which however can be performed easily and quickly. We will detail on this approach in the next two chapters.

A final aspect is the way in which the *information of the ecology is made available* to

the user. All systems discussed above focus on individual documents. The notes, annotations, hyperlinks and tags of users are displayed directly within the documents. However, no higher-level view is given that focuses on the relations that constitute the information ecology. We present an ecological view that fulfills this requirement.

In this chapter, we have reviewed the state of the art and identified open research questions. We will address these questions in the following chapters. In Chapter 4, we will introduce a theoretical interaction model of Pen-and-Paper User Interfaces that addresses the limitations of currently existing models. Next, in Chapter 5, we will present novel interaction techniques and visualizations and an integrated system framework for paper-based knowledge work. These are based on the theoretical interaction model.



# 4

## Interaction Model of Pen-and-Paper User Interfaces

In the two previous chapters, we analyzed how knowledge workers interact with documents and we reviewed the state of the art of paper-based user interfaces. We have identified significant shortcomings in existing models of Pen-and-Paper User Interfaces (PPUIs). In this chapter, we therefore contribute a theoretical interaction model of PPUIs.

The remainder of this chapter is structured as follows. We will first explain the concept of Pen-and-Paper User Interfaces in Section 4.1. We will then discuss three key dimensions of interaction in Pen-and-Paper User Interfaces. In Section 4.2, we present a model of the actual *interactions*. Based on an analytical separation of a semantic and a syntactic level of interaction, we identify interaction primitives that act as building blocks for performing semantic activities with a digital pen and paper. The model is of analytical value and provides support for the design of paper-based interfaces. The second dimension of the model concerns the *information* that is distributed across printed and digital media. We define document types, introduce two principles of how physical and digital information is related and provide a taxonomy of how printed and digital representations should be chosen to best complement each other. Finally, we model how users *collaborate* in a paper-based environment. This covers co-located and remote collaboration as well as privacy aspects. This analytical model was developed in an inductive empirical process and is grounded on our field work and on an analysis of existing user interfaces from related work. The model is the foundation of the interaction design of CoScribe, which will be presented in the next chapter.

It is helpful to define key terms before presenting the model. Our definition of an interaction model is based on the definition of [BL00]:

**Definition 5** (Interaction Model). *An interaction model is a set of principles, rules and properties that guide the design of an interface.*

Examples of interaction models are the WIMP model and Direct Manipulation [Shn83]. In contrast to implementation models, such as the Model-View-Controller concept [KP88], interaction models guide the design of the interface and of interaction techniques and not their implementation.

A widely accepted definition of interaction techniques is the following:

**Definition 6** (Interaction Technique [FvDFH90]). *An interaction technique is a way of using a physical input/output device to perform a generic task in a human-computer dialogue.*

**Definition 7** (Paper-based Knowledge Work). *Paper-based knowledge work is characterized by the concurrent or interleaving work with information contained on printed documents (such as books, articles, printouts of digital documents, handwritten notes) and with digital information on a computer.*

Today, paper-based knowledge work is still the most prevalent form of knowledge work. On the one hand, knowledge work which does not include the use of computers hardly exists any more. On the other hand, as shown in Chapter 2, paperless offices are far from reality, as paper offers unique affordances which are not provided by digital tools.

## 4.1 Pen-and-Paper User Interfaces (PPUIs)

**Post-WIMP user interfaces.** One way to overcome the rupture between the paper world and the digital world are Pen-and-Paper User Interfaces (PPUI). PPUIs are part of the larger class of user interfaces which is called post-desktop or post-WIMP user interfaces. These interfaces go beyond the desktop metaphor and diverge from the “window, icon, menu, pointing device” (WIMP) paradigm of classical Graphical User Interfaces. Van Dam defines post-WIMP interfaces as “containing at least one interaction technique not dependent on classical 2D widgets such as menus and icons” [vD97]. These interfaces build “on users’ pre-existing knowledge of the everyday, non-digital world to a much greater extent than before” [JGH<sup>+</sup>08].

Many post-WIMP interfaces extend interaction with computers into the physical space. Tangible User Interfaces, to which Pen-and-Paper User Interfaces belong, enable users to manipulate digital data by manipulating physical objects. In many cases, this results in a more direct and more natural interaction style. For example, physical cubes on a table might represent buildings. Moving a cube to another position automatically adapts the building’s position in the underlying digital CAD model. In mobile and ubiquitous computing, the interaction is not bound to a fix location. Users can freely move and interact at many places. Often, the user’s location in the physical space is taken into account and influences interaction and computation. Moreover, everyday physical objects, such as fridges, cups or chairs, are augmented by digital functionality.

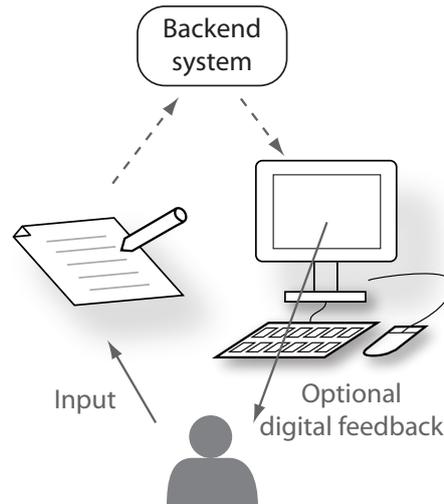


Figure 4.1: Basic setup of a Pen-and-Paper User Interface from a user perspective

**Basic setup of PPUIs.** PPUIs extend computing into the physical world by turning traditional paper into a digital interactive medium. As defined on page 43, Pen-and-Paper User Interfaces (PPUIs) rely on a digital pen that leaves visible ink traces on real paper sheets and moreover digitally tracks its movements on paper. This enables transferring digital handwritings and drawings to a digital system. Moreover, user interface elements can be printed onto paper which the user interacts with. For example, paper sheets may contain printed interface elements such as checkboxes, buttons, menus or fields for entering handwritten data or for issuing commands by drawing specific symbols. By interacting with a pen on these printed user interfaces, the user can control a digital system.

The basic setup of a PPUI is depicted in Figure 4.1. We distinguish two channels indicated by the arrows.

**Definition 8 (Channel).** *A channel is realized by a physical device and transfers data from the user to the digital system (input channel) or vice versa (output channel).*

In this basic setup, digital pen and paper serve as a pure input channel to feed data into the digital system. Optionally, the system provides feedback via a separate channel. This is typically realized by a nearby display of a computer, a mobile phone or a PDA, but the PPUI might also use auditory or tactile output channels. Hence, the PPUI may span several devices.

In the following chapters, we will see that CoScribe has a more complex PPUI. This enables several users to collaborate with multiple printed documents using a digital pen and printed tools. Updated versions of documents can be printed in order to support system feedback on paper. Moreover, the same digital pen as on

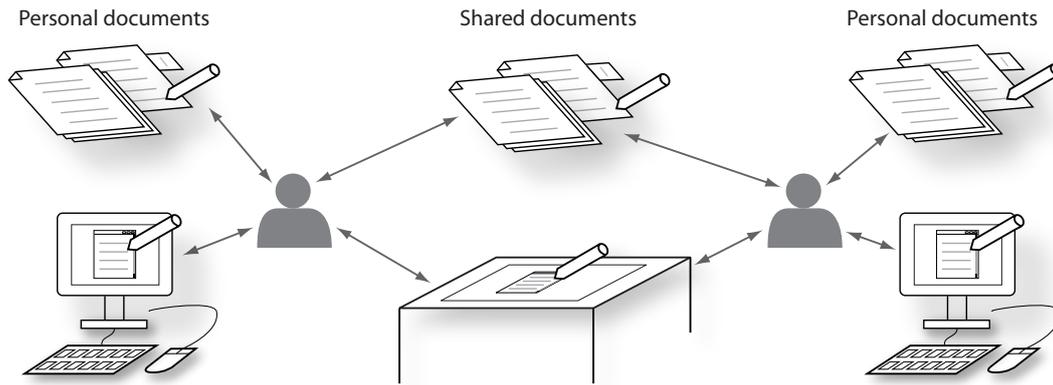


Figure 4.2: The extended setting of a Pen-and-Paper User Interfaces of our concept

paper is used to interact with digital documents on displays. This extended setup is illustrated in Fig. 4.2.

## 4.2 Interactions

A key question of our research is how to design pen-and-paper interactions that are intuitive, reliable and efficient. In this section, we present an analytical perspective on paper-based interaction which provides guidance for the design of paper-based user interfaces.

Pen-and-Paper User Interfaces have other characteristics than Graphical User Interfaces (GUI), as they comprise physical objects as interface elements and use a digital pen as the main interaction device. In contrast, GUIs typically rely on keyboards and mice. Moreover, paper is a very restricted output channel, which makes it challenging to design a user interface that supports complex activities and still remains simple to use. This implies that it is not sufficient to simply transfer GUI interactions to pen and paper. We argue that instead, interactions should build upon specific paper affordances, which GUIs cannot provide.

### 4.2.1 Semantic and Syntactic Levels of Interaction

Our approach for designing PPUIs is to let the user fully maintain her existing practices of working with a pen and printed documents. This accounts for the highly individual practices observed in contextual inquiries. However, while the system is able to store and display these informal artifacts made by the user, it cannot interpret them. If the user wants to inform the system about the semantics of the artifacts, she additionally performs more formal actions. These have to be simple, quick and reliable to be easily integrated into the work process without producing significant extraneous cognitive load.

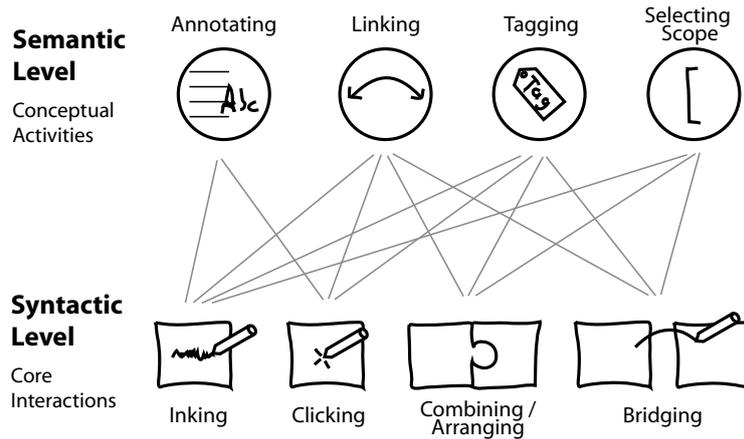


Figure 4.3: The analytical framework applied to CoScribe

The underlying principle of the model is an analytic separation into a semantic and a syntactic level of interaction (see Fig. 4.3):

1. The semantic level models *what* the user wants to do and comprises conceptual activities, i.e. the functionality offered by the user interface (for instance the activities of annotating, linking and tagging).
2. The syntactic level models *how* the user actually performs these activities. It comprises core interactions, i.e. primitive manipulations that are made with the PPUI in order to actually perform these conceptual activities (e.g. writing with the pen or attaching a paper sticker).

This separation of interaction into two levels is conceptually similar to the four-level model by Foley et al. [FvDFH90]. In addition to the semantic and syntactic levels, this covers a conceptual level, which is the user's mental model of the interactive system, and a lexical level, which encompasses the precise mechanisms by which the user specifies the syntax. We judge two levels to be sufficient for our aim of identifying core interactions and of modeling how these can be composed.

The challenge when designing a PPUI is first to identify simple and reliable core interactions which leverage the affordances of pen and paper. Second, the designer must decide which core interactions to use and how the user combines them to perform a conceptual activity. Ethnographic observations of users' current practices are an important method for informing these design decisions.

As we will see below in this section, this principle enables PPUIs that are both simple and complex: offering complex functionality while being intuitive and easy to use. This is mainly due to the fact that the generic core interactions are intuitive and easy, as they are inspired by the traditional cultural techniques of interacting with pen and paper.

### 4.2.2 Semantics: Conceptual Activities

The semantic level of interaction covers the functionality that is offered by the interface to support the user in his or her working tasks. Hence, the conceptual activities of a PPUI are not generic but depend on the purpose the PPUI is designed for.

For our application scenario, we identified three main conceptual activities that users perform when working with documents:

1. Extending documents by annotations
2. Creating references and hyperlinks to express relations between documents and passages of documents and
3. Creating tags in order to structure the problem domain by abstracting from given contents to higher-order concepts.

These conceptual activities are complemented by a fourth main activity, which is an existential part of the other activities:

4. Defining the scope of an annotation, link or tag. This is the portion of the document (or of several documents) it applies to. In the literature, the scope of an annotation is often referred to as its context. The scope of a link is called its anchor.

Conceptual activities can be hierarchically organized. For instance, scope selection is a sub-activity of the three other activities. Creating a hyperlink includes the sub-activity of selecting the passage of a document the hyperlink applies to. We might also define higher-level activities. For example, CoScribe supports the activity of “excerpting on a separate sheet of paper”, which relies both on annotation and on linking.

### 4.2.3 Syntax: Core Interactions

The syntactic level encompasses interaction primitives of PPUIs. These core interactions are independent of the functionality of the PPUI. Instead, they model in a generic way how users interact with pen and paper.

#### Empirical grounding

The core interactions are based on empirical evidence. We therefore briefly summarize the observations we made in an ethnographic field study on how users work with pens and printed documents. Following our observations, syntactic interactions with pen and paper can be divided into the following three categories:

1. *Writing*: Writing on an individual sheet of paper (e.g. annotations, handwritten links and keywords or symbols for tags). Participants often partitioned

the available space into separate functionality zones, e.g. reserving the left or right margins of the document for keywords in order to provide for a quick overview on all tags on a page.

2. *Moving, Arranging, Modifying*: Leveraging the material aspect of paper sheets, which can be flexibly arranged in the physical space. Specific spatial combinations of two or more paper sheets convey semantics (e.g. relating documents by putting them into a folder or on a stack or marking important pages with bookmark stickers). Moreover, the shape of physical paper sheets can be modified, e.g. by bending, folding and tearing.
3. *Pointing*: In the co-located collaboration of learning groups, we frequently observed pointing gestures to documents. Very often participants did not only point to one single document but consecutively to several documents or pages in order to express relations between the contents they pointed to.

### Core Interactions

These results lead to the following core interactions.

**Definition 9** (Core Interaction). *A core interaction is defined as an operation that a user performs by manipulating one or more page areas using a digital pen or her hands.*

Examples of page areas comprise a printed document page, a printed button element or an adhesive bookmarking sticker. First and foremost, page areas are contained on printed sheets of paper. However, by analogy, areas which are displayed on a display can equally act as page areas.

The operations performed with these page areas are the following:

- *Inking*: Writing with the digital pen on a page area. This includes free form handwritings and drawings that are digitally captured. Moreover, specific symbols and pen gestures may be performed to issue a command.
- *Clicking*: Performing one or more pen taps on a paper area to issue a command (e.g. on a printed “button” area). This is inspired by traditional pointing gestures. We distinguish clicking from inking for the following reason: While inking leaves visible pen traces and permanently alters the document, clicking is volatile, leaving it conceptually unchanged.
- *Moving*: Changing the physical location of the page area. This also includes picking it up and putting it down as well as flipping pages.
- *Altering Shape*: Altering the physical shape of a page area, for example by bending, folding or tearing it.

An important characteristic of paper is that it affords using several sheets of paper at a time. The following core interactions comprise the use of two or more sheets:

- *Combining*: Creating or modifying arrangements of page areas. This may be rather volatile (e.g. paper sheets laid out on a desk) or rather permanent (e.g. attached paper stickers, documents filed in a folder).
- *Bridging*: In contrast to physical combinations, bridging is a logical combination of several areas. This complements physical combinations or substitutes them when these are impractical or impossible. Inspired by consecutive pointing on several items, we model bridging as a connecting pen gesture on two areas.

The same core interaction may have a different meaning if it is performed on a different type of page area or using a different tool, e.g. a digital eraser.

## Discussion

We deliberately model combining and bridging as separate core interactions instead of combinations of inking, clicking and moving. The reason is that one of the most important affordances of paper is to support two-handed interaction and hence a really concurrent use of multiple sheets of paper.

Table 4.1 depicts how these core interactions are inspired by interactions with traditional paper and how they correspond to interactions within Graphical User Interfaces. As GUIs incorporate metaphors of traditional desks, they offer somewhat equivalent interactions for inking, clicking and moving. In contrast, traditional GUIs do not support two-handed input but only one single focus point. Consequentially, combining and bridging should be rather considered as sequential manipulations of individual areas. Moreover, the interaction of altering the shape leverages the physical specifics of paper environments and goes beyond what is possible in GUIs.

PPUIs should account for the rich interactions that are possible with paper. They should use a broad spectrum of core interactions. In Section 2.2, we have seen that paper affords the flexible spatial organization and the concurrent use of multiple sheets of paper. PPUIs should leverage this affordance and support the core interactions of combining and bridging. This stands in contrast to a design which is inspired by the GUI paradigm, in which interaction is restricted to single points of focus (due to the single mouse pointer).

Our interaction techniques to be presented in the following chapters rely on four core interactions. Instead of moving, we support combining, as cognitive activities typically consist of creating or modifying arrangements of several sheets but not of moving a single sheet without relation to others. Further, we very rarely observed manipulations of the physical shape of paper sheets in learning at universities. As moreover the tracking technology for detecting these manipulations reduces the mobility of paper use, this core interaction is not part of our design.

|                 | <b>Traditional Paper</b>                   | <b>PPUI</b>           | <b>GUI (following [BL00])</b> |
|-----------------|--|-----------------------|-------------------------------|
| Single sheet    | Writing                                    | <b>Inking</b>         | Text entry                    |
|                 | Pointing                                   | <b>Clicking</b>       | Pointing/Clicking             |
|                 | Moving                                     | <b>Moving</b>         | Dragging                      |
|                 | Altering shape (bending, folding, tearing) | <b>Altering shape</b> | - *                           |
| Multiple sheets | Arranging and Combining                    | <b>Combining</b>      | - *                           |
|                 | Subsequent pointing                        | <b>Bridging</b>       | - *                           |

\* No core interaction (performed by combining several core interactions)

Table 4.1: Comparison of Core Interactions

#### 4.2.4 Applicability of Core Interactions to Related Work

In this section, we apply the core interactions to a representative set of PPUI systems that support users in interacting with documents. We demonstrate that these can be classified in terms of the generic core interactions identified above. Table 4.2 provides an overview on these systems.

A first class of systems (e.g. [Gui03, WISN08, SN07]) augments paper documents by electronically capturing handwritten annotations (*inking*). ButterflyNet [YLK<sup>+</sup>06] additionally supports creating associations (*bridging*) between an area of a paper notebook and a digital photo with a pen gesture. Users can access the digital media by tapping with the pen on this paper area. PapierCraft [LGHH08] supports tagging paper documents with pen gestures (*inking*). In a calculator application, Wellner's Digital Desk [Wel93] supports entering numbers by pointing (*clicking*) on a number in an arbitrary document on the desk, regardless if it is printed or digitally projected. Moreover, the position of digital and physical documents on the desk can be mirrored to collaborators over distance in real-time. Each time a physical or digital document is *moved*, the position is updated at the remote site.

All these conceptual activities are performed with one single core interaction. However, there are also examples, where conceptual activities are supported by a combination of several core interactions. In PapierCraft, the user creates a hyperlink between two paper pages by first highlighting (*inking*) the passages that shall be linked and then *bridging* them with two consecutive markings on both pages. Moreover, PapierCraft supports creating physical collages: A user first physically *combines* two paper sheets by positioning one besides the other in a way that the margins slightly overlap. Drawing an associating line then digitally *bridges* both sheets. With the Digital Desk, users can select document snippets with a pen ges-

| System                             | Inking | Clicking | Moving | Altering shape | Combining | Associating | Purpose  |
|------------------------------------|--------|----------|--------|----------------|-----------|-------------|--|
| PADD [Gui03]                       | •      |          |        |                |           |             | Capturing handwriting  |
| PaperProof [WNS07]                 | •      |          |        |                |           |             | Capturing handwriting  |
| PaperPoint [SN07]                  | •      | •        |        |                |           |             | Capturing handwriting<br>Controlling slide presentations   |
| ButterflyNet [YLK <sup>+</sup> 06] | •      |          |        |                |           | •           | Capturing handwriting<br>Linking digital data  |
| DigitalDesk [Wel93]                | •      | •        | •      |                | •         |             | Drawing application<br>Selecting numbers<br>Moving shared documents<br>Crating physical/ digital collages                                      |
| PapierCraft [LGHH08]               | •      |          |        |                | •         | •           | Creating hyperlinks<br>Creating hyperlinks   |
| This thesis                        | •      | •        | •      |                | •         | •           | Capturing handwriting<br>Classifying annotations<br>Following hyperlinks<br>Controlling digital viewer<br>Creating hyperlinks<br>Creating tags |

Table 4.2: Core interactions of a representative set of related systems and of the interaction strategies presented in this thesis

ture (*inking*) and then *combine* these snippets in a physical and digital collage.

The novel interaction techniques presented in this thesis draw upon single core interactions as well as upon combinations of them. Table 4.2 includes an overview of these strategies, which will be discussed in the following chapters.

#### 4.2.5 Mapping Between Syntax and Semantics

Having defined the semantic and the syntactic level of interaction, we now focus on the intersection between both levels. We analyze how syntactic core interactions can be mapped to semantic activities, such as annotating or linking documents, and to their sub-activities. The latter are actions like specifying the first or the second link anchor, defining the scope of a document a link or tag applies to, defining a new tag category or applying a tag category to a document passage. The number of (sub-)activities is larger than the number of core interactions. For this reason, the model must rely on some kind of multiplexing between a smaller number of core interactions and a larger number of semantic activities.

In an influential interaction model, called Instrumental Interaction, Beaudouin-Lafon [BL00] discusses different types of multiplexing for tools in Graphical User

Interfaces. He analytically separates tools (which he calls instruments) into a physical and a logical part. While the physical part is the input device, for example the mouse, the logical part is realized in software on a screen, for example the button in a toolbar palette for selecting the instrument for drawing lines. He discusses two types of multiplexing between a single device and many logical instruments: temporal and spatial multiplexing. Temporal multiplexing means that the user selects an instrument which remains active until another instrument is selected. This corresponds to modes. Spatial multiplexing means that the instrument becomes active if the user positions the input device in an area of the screen which is associated with this instrument.

We adapt and extend Beaudouin-Lafon's types of multiplexing and distinguish the following five types of multiplexing, which are used to map core interactions to activities. These types do not exclusively apply to PPUIs, but are all possible in GUIs as well. A syntactic interaction that performs a specific activity might use one or more of these types of multiplexing between core interactions and semantic activities.

**Compositional multiplexing.** A key finding of our field work was that in traditional paper practice, users often combine several core interactions to perform a single conceptual activity. The user may for example attach an adhesive indexing sticker onto a document page and then write a label on this sticker. We call this compositional multiplexing and judge it to be – together with spatial multiplexing – most important for interaction in PPUIs. This principle enables users to draw upon a small repository of simple core interactions. These act as flexible building-blocks which are used and reused for multiple conceptual activities.

Figure 4.3 on p. 91 depicts how the model is used for the interaction techniques of CoScribe, which will be presented in the following chapter. CoScribe supports the conceptual activities of annotating, linking and tagging documents as well as of defining the scope of annotations, links and tags. The edges between the upper and the lower level indicate how the semantic activities of our application domain are composed of several simple core interactions.

**Spatial multiplexing.** In traditional practice of working with paper documents, different sheets of paper or different areas on one sheet have different functions. Actions can have different meanings depending on the type of paper medium which is used. For example, writing a label on an index sticker possibly has another meaning than writing the same term on a document page. Spatial multiplexing fits well with working with objects in the physical space. The first reason is that much space is available for different page areas, typically more than on a screen. Second, it is easy and intuitive to select a given area by physically grasping it.

Applied to our model of PPUIs this means that PPUIs should comprise different types of page areas that serve different purposes (e.g. document pages, folders and index stickers). The same core interaction performed on a different type of page area

leads to another result. For example, if the core interaction of inking is performed on an adhesive bookmarking sticker, it might define the label of an index, whereas inking on a document page might create a free-form annotation.

**Gestural multiplexing.** Different words or gestures written with the same pen on the same type of page area serve different purposes. For example the keyword “important” made in a margin column might be a tag while “cf. page 3” might be a link.

Gestural multiplexing is very powerful, as a large number of gestures can be defined. Moreover, it does not require other tools than one pen and unspecific sheets of paper. However, a heavy use of gestural multiplexing runs the risk to create a command-based interface which requires that the user memorizes a large number of possible counterintuitive commands instead of naturally working with documents. For this reason, CoScribe uses only a small set of very simple gestures, such as points, lines and tapping followed by holding down the pen, and puts the emphasis rather on a rich variety of paper tools.

**Temporal multiplexing.** The meaning of an action can depend on which action the user has made beforehand. This means that the system has modes. For example, several differently colored buttons for choosing ink colors might be printed on a page. By tapping with the pen on a button, the user selects the corresponding color. All subsequent inking interactions use this color until another one is selected.

Temporal multiplexing is critical, as this requires that the user performs an explicit action to activate the mode and moreover the user may become unaware of the currently activated mode. The latter issue is particularly crucial if in paper-only environments no real-time feedback on the current mode of the system can be given.

For this reason, the design of CoScribe avoids temporal multiplexing. Only the core interaction for bridging is modal, enabling the user to create two separate end points one after another. But even this interaction exists in a modeless variant where both end points are defined by one single action.

**Device multiplexing.** If performed with another device, the same actions can have a different signification. For example, inking with a digital pen which has a ball-point tip might be used for making textual annotations, whereas inking with another type of digital pen having a highlighter tip might be used for marking passages. Device multiplexing is powerful if the number of devices keeps manageable and if the repartition of functions across devices is clear and intuitive.

CoScribe does not use device multiplexing because one of our goals is to reduce the costs of devices to make digital pen-and-paper interaction not significantly more expensive for the end-user than traditional pen and paper. Moreover, there is empirical evidence that students tend to use one single pen rather than switching between many tools [Mar97].

## 4.3 Information

This section models *information* and specifically how it is distributed across physical and digital media.

In Chapter 2, we have seen that knowledge workers interweave the use of information contained in printed and digital documents. Paper-based knowledge work is characterized by a gap between printed and digital media. This is called *media discontinuity*. Integrating information from both worlds requires the knowledge worker to perform manual workarounds, all incorporating a significant additional effort. Some users print all relevant digital documents to dispose of all information in the same medium. Other users tend to digitize all relevant information as quickly as possible, for example by making typewritten excerpts from books. Another workaround consists of expressing relationships between printed and digital information by references that are manually added to the documents (e.g. “see Wikipedia: Quicksort, paragraph on complexity”). While this approach does not require transforming documents to another representation, it is complicated and may lead to inconsistent references. In contrast, PPUs support a closer coupling of printed and digital information.

In this section, we first present different types of documents and how these are represented. We will then discuss two principles that model how information is distributed across physical and digital representations. Finally, we will derive abstract design principles.

### 4.3.1 Document Types and Representations

Information is mainly contained in documents, as the model aims at supporting knowledge workers in interacting with documents. But the user also deals with information which is not contained in documents, such as feedback generated by the system on success or failure of the user’s actions and feedback on the current state of the system.

Based on the three dimension of documents identified in [Péd03], which are *form*, *sign* and *medium*, we define a document as follows:

**Definition 10** (Document). *A document is made up of content which is materialized on a physical or digital medium in a stable structure and which conveys semantic meaning.*

Documents and other contents can exist in different representations:

**Definition 11** (Representation). *A representation is the materialization of a document or of other contents on a specific medium.*

For instance, the same contents might exist in a representation which is printed on paper and in a digital representation on a screen. Note that representations do not only apply to documents, but also to other contents commonly not characterized as documents, such as system feedback.

Based on the media that can be used to represent the contents of a document, we distinguish three types of documents, all used by knowledge workers:

1. *Physical-only documents* are available in printed form but the user does not dispose of an equivalent digital counterpart. The corresponding digital representation does not contain the contents of the document but only some metadata like the title, the names of the authors, the year of publication, an image of the book cover and so on. Examples of this class include printed books and journals.

CoScribe provides restricted support for this type of documents. Using a specific paper sticker attached onto the cover of the book or journal, the user can utilize the digital pen to interact with the *entire* document, but it is not possible to digitally address or annotate subparts of the document (see Section 5.4.2 for more information).

2. *Hybrid documents* have similar representations in printed and in digital forms. Their digital contents can be transformed to a spatially fix, temporally static and two-dimensional layout on one single or on several separate sheets of paper. No essential contents are lost and an unambiguous mapping from the printed to the underlying digital contents is possible. This is particularly the case for digital document types which are specifically designed for printing purposes like Adobe's Portable Document Format or Microsoft's XPS Format, but also for many other documents, such as word-processing documents, presentation slides, images and Web pages which do not contain time-continuous contents.

CoScribe provides most comprehensive support for hybrid documents.

3. *Digital-only documents* are available in a digital form but cannot be represented on a purely physical medium without losing the essence of the information. As a consequence, digital-only documents either do not have a physical representation or they have one which however does not contain the information of the document but only acts as a proxy to access the digital information from the physical world. This includes time-continuous media such as audio and video documents and digital animations, but also interactive documents such as electronic quizzes.

The current prototype of CoScribe does not support digital-only documents. However, it could be easily extended to support audio or video recordings which are automatically linked to handwritten annotations, links and tags (similarly to the work of Stifelman et al. [SAS01]).

### 4.3.2 Relations between Physical and Digital Information

Having discussed different types of documents and their representations, we will now move on our focus to the media discontinuity and address how information

|                      | EQUIVALENCE   | COMPLEMENTARITY   |
|----------------------|---|---|
| Piece of Information | <p><b>single</b></p> <p>one single piece of information</p> | <p><b>multiple</b></p> <p>several complementing pieces of information</p>           |
| Representation       | <p><b>multiple</b></p> <p>multiple representations</p>      | <p><b>single</b></p> <p>one single representation for each piece of information</p> |

Figure 4.4: Distinctive features of the two principles

is distributed across printed and digital representations and how these are related. Not only documents exist in different representations, but also other information, such as system feedback. For this reason, we use the more general term *piece of information*.

We distinguish two orthogonal principles that model the relations between printed and digital pieces of information on a semantic level: equivalence and complementarity. We discuss each principle in turn.

### Principle 1: Equivalence of Information

We refer to the first principle as *equivalence*. It is illustrated in Figure 4.5 (left). This principle captures that the same piece of information can be transformed between equivalent representations on different media. The user can choose the representation which best fits her needs. As a matter of course, this is only possible for hybrid documents, but not for printed-only or digital-only documents, which do not have equivalent representations in the other world.

On the level of information, both representations are equivalent, as they contain the same contents. However, different representations can offer different affordances. For example, the user may utilize a printer to transform a digital representation to a printed one because it is more convenient to read and annotate information on paper. In the reverse direction, she may transform a printed piece of information to an equivalent digital representation that better affords searching for specific terms.

Two or more equivalent representations of the same piece of information can be used exclusively or in combination at the same point of time. The latter combines the affordances of several representational media. For example the printed representation of a document can be used to quickly navigate between different pages

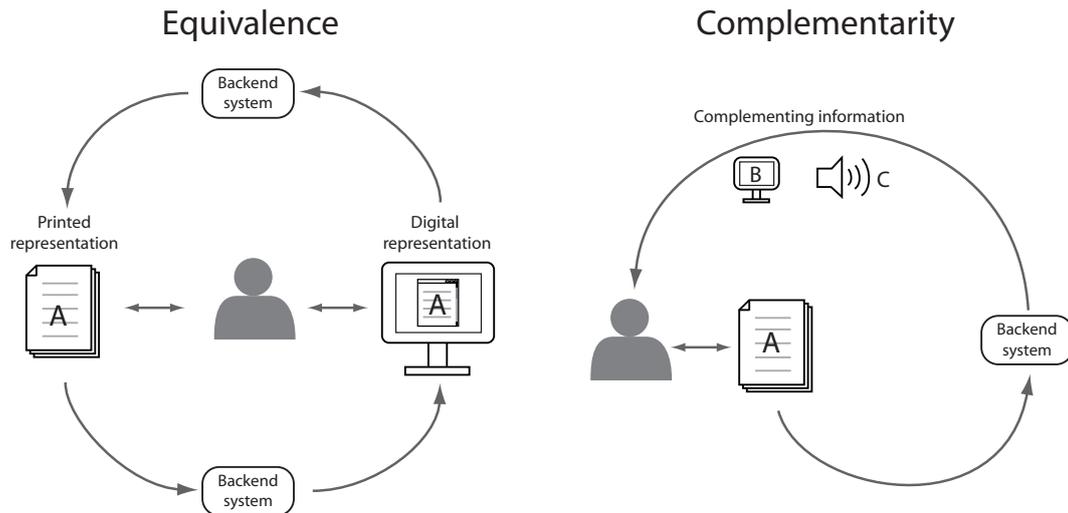


Figure 4.5: The two principles of how information is related between different representations

of this document, while the digital representation better affords editing, moving or deleting existing annotations, as it is updated in real-time.

### Principle 2: Complementarity of Information

The second principle, called *complementarity*, models the orthogonal dimension. It is illustrated in Figure 4.5 (right). The entire information consists of several, different pieces of information that are distributed across different representations and complement each other. Each piece of information is available in one fix representation. The distinctive features of both principles are shown in a matrix in Figure 4.4. In practice, both dimensions appear together. For example, one of the complementing pieces of information can be transformed to an equivalent representation in turn.

Complementing information serves two main purposes. First, complementing pieces of information can provide the user with *updates* of a piece of information. While it is usually preferable to directly update the representation of the piece of information the user is working with, this is not always possible. For example, a digital system cannot update information on paper in real-time. However, real-time updates are possible with a complementing digital representation.

Second, complementing information can provide *different pieces of information* on a second channel. For example, if the user works with a printed representation of a Web page which contains background music, a complementing digital channel may play this music. A second example is system feedback, which may be given on another channel, i.e. in another representation.

### Research on Multimodality

In the field of multimodality, a considerable body of research analyzes how several modalities can be combined. We will now put our principles in this context.

In contrast to research on multimodality, which focuses on interaction and on input and output channels, our principles model the relations between information on a semantic level. Moreover, we do not attempt to provide a general characterization but specifically focus on the relations between paper and complementing digital channels.

The CARE properties [CNS<sup>+</sup>95] define four principles of how several modalities are related: complementarity, assignment, redundancy and equivalence. Our principle of equivalence encompasses both the CARE principles of equivalence and of redundancy. First, the user has the choice between several modalities, which are equivalent. Second, she can either exclusively use one single of them or combine several modalities (redundancy). Our second principle of complementarity corresponds to the equally entitled CARE principle. The fourth CARE principle – assignment – is not reflected in our principles, as assignment means that only one single modality can be used and we model relations between printed and digital representations.

Vernier and Nigay [VN00] developed a framework for the combination of output modalities that consists of several levels. Its semantic level defines complementarity and redundancy. This corresponds to our two principles. Moreover, the semantic level comprises a third principle for information that is concurrently used but that has no related meaning. In our model, this stands implicitly behind the two principles. The temporal and spatial levels of the framework correspond to our classification of feedback channels.

This brief review demonstrates that our principles can be categorized using taxonomies from related research.

#### 4.3.3 Appropriate Complementing Representations

We now analyze which complementing representations should be chosen. Table 4.3 provides an overview on combinations of representations and their properties. We examine complementing representations for information that is represented on paper and on a display. Complementing representations are visual, auditory or haptic. Combinations of representations can be classified by three dimensions:

1. The time needed for updating information on the complementing representation. Most often, system feedback on user input should be provided immediately.
2. The spatial distance between the complementing representations. They should be located as near as possible.
3. The amount of data that can be transferred using the complementing representation. The data rate should be high to provide comprehensive informa-

| Representation | Complementing representation | immediate           | near | high data rate |     |
|----------------|------------------------------|---------------------|------|----------------|-----|
| Paper          | Visual                       | Overlaid projection | •    | •              | •   |
|                |                              | Nearby display      | •    | –              | •   |
|                |                              | Re-print            | –    | •              | •   |
|                |                              | Pen display / LED   | •    | •              | –   |
|                | Auditory                     | Speaker             | •    | •              | •/– |
| Display        | Haptic                       | Pen                 | •    | •              | –   |
|                | Visual                       | Display             | •    | •              | •   |
|                | Auditory                     | Speaker             | •    | •              | •/– |
|                | Haptic                       | Pen                 | •    | •              | –   |

Table 4.3: Appropriate complementing representations

tion (e.g. the contents of a document page), while representations with lower data rates can only be used for purposes such as indicating success or failure of an operation or the current mode of the system.

Providing feedback for *input made on a digital display* is not critical. Extensive feedback is best provided on the same display or on a separate auditory channel. Less extensive feedback can also be provided by the digital pen if supported by the given digital pen technology. All these channels are both near and immediate.

In contrast, near and immediate extensive feedback for *input made on paper* is only available through auditory channels. However, while this channel offers a high data rate for conveying textual information, it cannot be used for transferring visual information. Near and immediate visual feedback with high data rates requires overlaid projection and a complex technical setup with a camera and projector mounted above the interaction surface. If this restriction is not acceptable, extensive visual feedback can only be achieved by accepting trade-offs with respect to distance or immediacy. On the one hand, immediate feedback can be provided on a nearby display, which however is not co-located with the input medium. On the other hand, an updated printout of the paper representation can contain comprehensive updates and feedback, which is co-located, but which is necessarily delayed.

CoScribe uses the following complementing representations for printed representations: If it is not appropriate to print an updated version of the representation, additional information is provided on a nearby display. System feedback on its status and on success or failure of actions is performed by more restricted channels, which however have a smaller spatial offset with respect to the input. These channels are audio as well as visual and haptic feedback provided by the pen.

|              |              | Same Time                      | Different Times                        |
|--------------|--------------|--------------------------------|--|
| Diff. Places | Same Place   | <b>Co-located interaction</b>  | Asynchronous local interaction         |
|              | Diff. Places | Synchronous remote interaction | <b>Asynchronous remote interaction</b> |

Figure 4.6: Two dimensions for classifying collaborative user interfaces. While all four resulting types are supported by CoScribe, its specific focus is on the types in bold typesetting.

## 4.4 Collaboration

Learning at universities comprises various forms of collaboration including co-located and remote as well as synchronous and asynchronous collaboration. In the previous sections, the model accounted only for single users who interact with the Pen-and-Paper User Interface. As our goal is to support collaboration around paper-based and digital documents, we now add the dimension of multiple persons using the system.

The traditional way to classify collaborative user interfaces is by a matrix of time and space [EGR91] (see Fig. 4.6). Collaborating users may work together at the same time, engaging in a synchronous exchange, or they may work at different points in time, leaving messages that collaborators access later on. On the second dimension, collaborating users may be located at the same physical place (e.g. a meeting room) or may be physically distributed (e.g. each person being at her home) and collaborate over distance. In the following, we will discuss each of the types taking on a paper-centric perspective.

### 4.4.1 Local Collaboration

Co-located collaboration occurs frequently in knowledge work, for example in business or learning meetings and seminars. We discuss why physical interaction is well-suited for co-located collaboration and in which ways documents are collaboratively used.

**Multiple points of interaction.** Multiple physical objects in the physical space serve co-located interaction better than an ordinary computer setup, as the latter provides

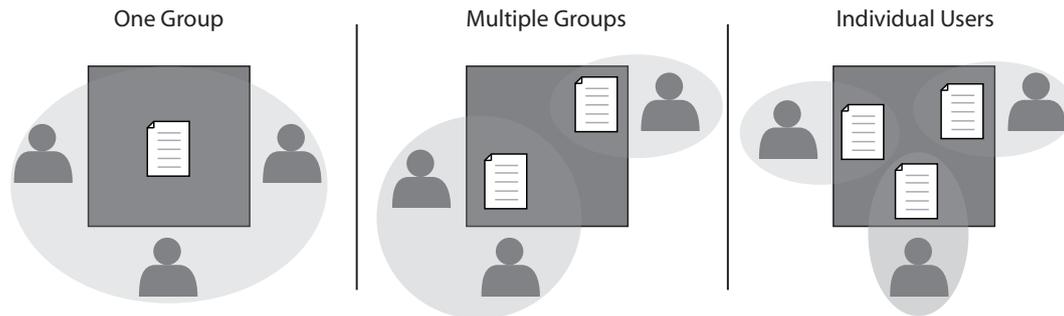


Figure 4.7: Multiple points of interaction enable various co-located collaborative settings

only one single point of interaction. This includes both input and output channels. Further, it disposes of one single set of input devices (typically a mouse and a keyboard). For this reason, several users cannot operate the system concurrently but only in an alternating manner and it is not easy to keep track which person performed which activity. Moreover, the restricted screen size of ordinary computers makes it difficult to display a larger number of documents at a time.

In contrast to this digital setup, interaction in the physical space provides for multiple points of interaction. In a collaborative PPUI, each user disposes of his or her personal pen. Multiple pens can be used simultaneously, even on one single physical sheet of paper or on one display. Hence, the concurrent use by multiple persons is supported and activities can be attributed to individual users. Moreover, a large number of display surfaces is available (considering paper documents as a kind of display), which can be easily repositioned and combined. This provides for very flexible collaborative settings. Figure 4.7 depicts three example settings.

**Synchronous and asynchronous document use.** As co-located interaction is classified in the *same time* column of the matrix, one might think that co-located interaction means that several users work concurrently on the same document. However, in an ethnographic study we observed that participants in meetings do not only jointly work on shared documents, but there are also phases in which each user individually works on individual documents (see Fig. 4.7 right). After having finished working with a document, he or she may pass it to a collaborator, who then in turn works individually on it. If the medium the document is represented on is mobile (e.g. paper, a laptop, a PDA or a small display), passing is typically done by physically handing over this medium to the other person. If the medium is not mobile (e.g. a large screen), users change their places. From a document point of view, this form of collaboration is asynchronous, although several users work at the same time at the same place. We thus distinguish *synchronous co-located* from *asynchronous co-located* collaboration. In typical settings, both forms occur and are interleaved.

Besides co-located interaction, users can collaborate at the same place in different points of time, i.e. without meeting each other. This form is called *asynchronous local* collaboration. This is for example the case for classic blackboards or group calendars. At a given location, users can leave written messages for subsequent users and can view messages left by other users earlier on. This type of collaboration is supported by PPUIs as well. A user can leave printed documents at a given location for subsequent users. Similarly, he or she can work on digital documents using a local pen-enabled display. Later on, subsequent persons can access these document using the same display.

#### 4.4.2 Remote Collaboration

While paper is a very adequate medium *during* co-located meetings, it is hard to make available the information which is contained on printed documents to all participants *after* the meeting. Typical workarounds consist of photocopying relevant documents or of asking one participant to typewrite the information and to share the resulting digital document by e-mail. Yet, comments made on personal documents which are judged less important are typically not made available to other participants. It is even more complicated to share with fellow students the annotations which were made on handouts during lectures. This requires photocopying a large number of pages or manually transcribing important notes from the scripts of fellows. As a consequence, lecture notes mostly remain private.

PPUIs can address this challenge and enables users to collaborate over distance not only with *digital* but also with *physical* documents. The collaborating users are working at different locations and possibly at different points in time. Each user works on her local physical or digital representation of the document. His or her changes are automatically distributed to collaborating users over a network connection. These can then display a local representation of this document on the screen or can print it on paper.

Remote collaboration does not only apply to settings where users are working at different locations at the time, but it is also an important element that complements co-located collaboration. After the co-located phase, all users can access digital versions of shared documents over distance, regardless if a shared document was used in a physical or in a digital representation. If it is a hybrid document, they can even print out a personal copy.

**Synchronous collaboration.** In addition to the asynchronous support, PPUIs can also be used for remote synchronous collaboration. Several spatially distributed users can work with the same document page at the same point in time, for example writing on it, creating links or tags. This page acts as a shared workspace. All activities are transmitted to remote users in real-time and can be displayed on a screen as complementing information.

### 4.4.3 Privacy

An important aspect about the collaborative use of documents is privacy. It must be ensured that personal contents are not accessed by other users. When modeling privacy, it is crucial to account both for social protocols and for technical solutions that allow to define explicit access rights.

**Personal and public spaces.** In co-located collaboration, the management of access rights typically does not require explicit technical mechanisms. Instead, social protocols are used to define which documents of other persons one has the right to access. In this context, space and location are influential factors. Locating a document in a personal space implicitly conveys that the owner does not want other collaborators to access this document. Correspondingly, public spaces are used for documents that are shared with other users. For instance, a printed document which is located in front of a user could be private while a document located in the middle of the table in a similar distance from all users could be public. Documents on a personal laptop are typically private while documents that are displayed on a shared screen are public. PPIs support well these social protocols because they heavily draw upon the interaction in the physical space and include the use of several display surfaces made up of paper and/or devices.

**Digital access rights.** A purely implicit management of access rights is less appropriate for remote interaction. In co-located settings, the user can control if the social protocol is followed by all collaborators. For instance, if a collaborator would not adhere to the protocol and access documents on the personal laptop of another user, one would become aware of this violation. In contrast, this kind of social control is not possible in remote collaboration unless specific awareness functionality is provided that informs which user has accessed which of his or her personal documents. Even if this awareness is provided, social control might not work if the setting includes a large number of participants, many of them being strangers to a given user. For these reasons, remote sharing settings need explicit mechanisms for defining the rights for accessing digital contents. These explicit mechanisms are “hard” in the sense that the computer system prevents access to contents which one has not access privileges for.

## 4.5 Conclusions and Design Guidelines

In this chapter, we presented an interaction model for Pen-and-Paper User Interfaces (PPIs). The model provides a conceptual basis for the design of paper-based interaction techniques and systems. It was developed in an inductive empirical process and is grounded on our field work and on an analysis of existing user interfaces from related work. The model defines Pen-and-Paper User Interfaces and covers three key dimensions of PPIs in knowledge work: interactions, information and collaboration.

This model implies the following design guidelines for Pen-and-Paper User Interfaces:

**1. Draw upon the affordances of paper and upon traditional practices of working with paper instead of mimicking GUIs**

Traditional practices of working with paper have evolved over a long period of time and have proven to be highly efficient in many situations. As they leverage the affordances of paper, they have advantages that GUI-inspired interfaces do not necessarily provide (e.g. traditional interaction is reliable even though there is no real-time feedback provided by some kind of system). In order to inspire the design of a paper-based application, the designer should use *ethnographic methods* to analyze what activities the system should support (*semantic level* of the interaction model) and by which means they actually perform this activities (*syntactic level*).

**2. Use simple and flexible building blocks**

The interaction model proposes a *small inventory of generic core interactions*. Each core interaction is inspired from traditional practices of interacting with paper and designed to be intuitive, simple to use and reliable. Using *compositional multiplexing*, core interactions can act as building blocks and can be flexibly combined. This provides for offering complex functionality in a PPUI which nevertheless remains easy to use, as only a very restricted number of core interactions is used on the syntactic level.

**3. Provide for rich interactions**

The design should incorporate the richness of interacting with paper. First, this includes using a *wide variety of core interactions*, in particular those that combine multiple sheets (e.g. combining/arranging and bridging). This stands in contrast to a design that would be inspired from the single point of focus of Graphical User Interfaces and that would merely build upon the core interactions of inking and clicking. Second, PPUIs should include *tangible tools* that are made out of paper (e.g. leaf binders, index stickers). This uses *spatial multiplexing*, binding different functionality to different sheets of paper and thereby to different positions in the physical space. This results in more intuitive interfaces that moreover can be flexibly adapted to the current work situation by arranging tangible tools accordingly.

**4. Cope with the restricted feedback capabilities of paper**

A major challenge for the design of PPUIs are the very restricted feedback capabilities that we have when using paper as an interactive medium. First and foremost, the designer should identify interactions that leverage *feedback which is provided by the physical properties of paper* instead of digital real-time feedback which must be provided by the system. For instance, writing on a sheet of paper generates “real-time feedback” by leaving visible ink traces. Attaching

an index sticker generates feedback in form of a sticker which is visible on paper. Second, the design should *avoid temporal modes* if the current mode cannot be clearly and continuously communicated to the user. Instead, the design should leverage spatial multiplexing or device multiplexing – this provides for “modes” which the user can clearly recognize as such without digital feedback. Third, the design should make *careful use of handwriting recognition and gesture recognition*. These should be used as a central part of the interaction design only if the recognition does not imply uncertainty or if real-time feedback on the recognition result can be given to the user. Finally, if digital real-time feedback should be provided, the designer should *choose an appropriate channel* following the three-dimensional taxonomy of our model (temporal/spatial distance, data rate).

#### 5. Leave interactional freedom to the user

As the practices of working with pen and paper are highly informal and individual, PPIs should leave much freedom to the user and impose only minimal constraints on traditional, well-established practices.

#### 6. Account for paper-based privacy management

An important issue of collaboration around documents is privacy. We point out that for an effective privacy management, it is not sufficient to only consider technological solutions, but that instead implicit social protocols are a key part of the practices of information ecologies. In co-located settings, privacy is typically ensured by implicit social protocols. These mainly rely on private and public spaces where documents are located. In contrast, remote collaborative settings require more explicit digital access rights ensured by a technological solution. The design should take into account both forms of privacy management.

In the next chapter, we will integrate the findings presented in the first four chapters of this thesis, moving on from the analytical and model-centered point of view to a design perspective. We will thereby take up the requirements that we have identified in Chapter 2 and that we have found in Chapter 3 to be not addressed by previous research. We will apply the theoretical model, which was presented in this chapter, and contribute novel interaction techniques and a system framework for paper-based knowledge work, which address these requirements.

# 5

## CoScribe: Collaborative Cross-media Annotation, Linking and Tagging

In the previous chapter, we have introduced an interaction model for collaborative Pen-and-Paper User Interfaces. This model is the foundation for CoScribe. CoScribe is an interaction concept and system that supports knowledge workers in reading, relating and integrating the increasing amount of information one has to deal with. Scientists and engineers have early thought of this problem domain, generating such influential visions as those of Vannevar Bush [Bus45] and Ted Nelson [Nel65]. CoScribe addresses this challenge from a paper-centric perspective. It includes novel interaction techniques that tightly integrate the physical world of printed documents and the world of digital documents.

CoScribe takes on an integrated, ecology-centered viewpoint. As discussed in Section 2.4, an information ecology comprises various practices of working with documents, the combined use of printed and digital artifacts and various forms of collaboration with other persons. CoScribe focuses on the interdependencies between users, physical and digital artifacts and practices in a given work setting. Whereas previous solutions supported only a small fraction of the practices which are frequent in typical information ecologies, CoScribe provides more comprehensive functionality: It supports three central and generic activities of working with documents (annotation, linking and tagging). Moreover, enables both co-located and remote asynchronous collaboration around the rather static medium of paper. Further, CoScribe simplifies pen-based interaction, as the same device can be used for all activities both on paper and on pen-enabled displays. Finally, CoScribe creates a richer user experience than previous work by offering a wide set of intuitive tools which are made out of paper. Figure 5.1 gives an overview of CoScribe.

The remainder of this chapter is structured as follows. First, in Section 5.1, an overview of the entire concept and of the underlying assumptions is given. In Section 5.2, we then address by which means CoScribe integrates printed and digital

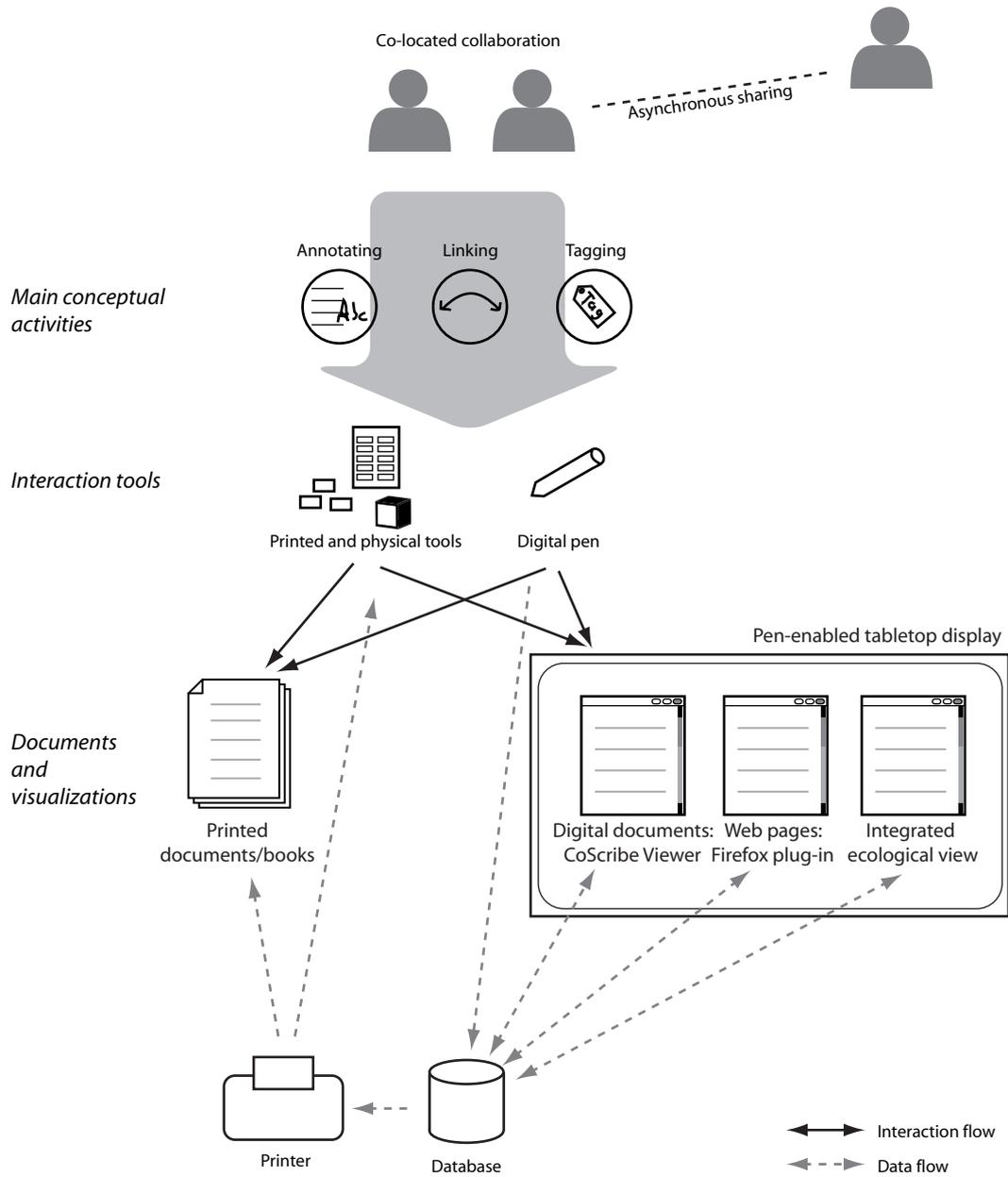


Figure 5.1: Overview of CoScribe

documents. This presents the concept of unified interaction using the same digital pen on paper and on interactive displays. It moreover shows how printed documents and tools are used in CoScribe, focuses on the digital representations of documents and details on means that enable users to switch between printed and digital document representations. In the subsequent three sections, we discuss the novel interaction techniques and visualizations of CoScribe in more detail. In Section 5.3, the interaction concepts and visualizations for collaborative handwritten annotations are introduced. We contribute the concept of user-adaptable printed user interfaces and present mechanisms for the paper-based sharing of annotations and for their collaborative visualization. In Section 5.4, we present an interaction technique for creating and following cross-media hyperlinks. These closely integrate paper and digital media, as they span both types of documents in any combination and are created on both media with one single pen-based association gesture. The concept of specific association areas provides for very flexible link scopes while the interaction remains intuitive, quick and easy. Like annotations, hyperlinks can be shared with other users. Finally, in Section 5.5, we contribute four novel pen-based and tangible interaction techniques that provide for tagging documents and processes. These are based upon a wide variety of interaction metaphors, mainly inspired from the rich practices of using traditional paper.

## 5.1 Overview of CoScribe

### 5.1.1 Pen-based Interaction with Printed and Digital Documents

While knowledge workers typically work both with printed and digital documents, a large gap remains between both worlds and the interactions remain fragmented. The user must deal with different interaction devices and techniques and it is difficult to digitize information which is contained on physical paper.

In contrast, CoScribe aims at a very seamless integration of physical and digital documents. It offers the same tools and interactions for both types of documents. Moreover the user can choose between working with a printed representation of a document and with a digital representation, depending on his or her preferences and the current context.

Users can print digital documents onto real paper and interact with them using a digital Anoto pen [Anob] (Fig. 5.2). This digital pen can be used like a traditional pen. It leaves visible ink traces on paper and has a similar form factor. The user can naturally switch between different sheets of paper without any additional action being necessary (e.g. pushing a button for indicating page changes or calibrating the new sheet of paper). Even if the digital functionality is not working for some reason, a physical copy of the handwritings is created (graceful degradation). In addition, the pen captures digital ink traces by continuously tracking both on which paper sheet it is located on and its position on that sheet.

Changes made on a printed document are automatically included in its digital version and available in the CoScribe viewer (Fig. 5.3). In addition to printed doc-



Figure 5.2: Anoto pen, model Nokia SU-1B (Photo: © Nokia)

uments, CoScribe offers various specialized tools made of paper (such as folders and index stickers). Tools and documents can be easily printed by the user on an ordinary printer using the CoScribe print toolkit. As the main functionality only requires a digital pen and paper, CoScribe supports mobile use and allows for instance to read and annotate documents in the lecture hall, at home or in public transport. The same digital pen can also be used on an interactive table to interact with digital documents. This will be detailed below.

CoScribe was designed with the goal of being intuitive and simple to use although it offers a large number of functionalities. As we will see, the pen-based interaction techniques are easy to learn and to integrate into current practices of working with documents. For this reason, our approach is to let the user fully maintain her existing practices of working with a pen and printed documents. This accounts for the highly individual practices observed in contextual inquiries. However, while the system is able to store and display these informal artifacts made by the user, it cannot interpret them. If the user wants to inform the system about the semantics of the artifacts (e.g. “this is a reference”, “this is a tag”), she additionally performs more formal actions. These are simple, quick and reliable in order to avoid producing significant extraneous cognitive load. We will detail on these interactions below.

### 5.1.2 A Toolset for Document-based Knowledge Work

CoScribe offers a set of generic activities of working with documents: annotating, linking and tagging documents. These activities enable the knowledge worker to read and understand documents and then to relate and abstract them in order to gain structural knowledge of the problem domain. Currently supported document types are PDF and PowerPoint documents, Web pages and physical books.

As CoScribe supports rather generic activities, it is not tailored to one specific



Figure 5.3: The CoScribe viewer offers access to digital versions of printed documents

purpose. Instead, it can be considered as a toolset for document-based knowledge work. The knowledge worker chooses appropriate tools, combines them and uses them in a way that best fits his or her current task. This leaves much freedom to the end-user. The flexibility of CoScribe is best illustrated by giving examples of settings it can be used in. In our application scenario of learning at universities, this includes taking notes and making annotations during courses, reviewing own notes and shared notes of other learners, preparing for exams in learning group meetings, excerpting documents, searching and integrating literature for preparing an article or a term paper and even giving presentations by controlling the slide actually being presented using a printout. Despite its wide applicability, CoScribe remains easy to use, as it relies on a small set of simple, but generic interaction strategies, which are inspired from traditional practices of working with paper documents.

Throughout this chapter, we illustrate the interaction concepts of CoScribe using a real-world example scenario:

**Scenario 1.** *Sally is a second-year university student and attends several lectures and seminars. Her work is mainly based on documents, such as physical books, pages with her handwritten notes, printouts of web pages, digital PDF articles and digital web pages. While she often deeply engages with one single document, for example for reading and understanding a complicated scientific article, an important part of her work also consists in integrating contents from various sources to understand an issue. She often works on her own, but is also frequently involved in collaborative settings, for example during the sessions of seminars or in meetings with fellow students. She has various workplaces, working at home, in lecture halls, at the library and in public transport. Hence, both the way Sally works with documents and the contexts of this work are highly varied. CoScribe supports all these forms of document work. In the following Scenario boxes, this is discussed in more detail.*

### 5.1.3 Co-located and Remote Collaboration

The ecological viewpoint of CoScribe also focuses on the relations between multiple users in a given work practice. While the functionality of CoScribe is helpful for individual users, its strength lies in supporting collaboration of various users.

**Scenario 2 (Learning Group Meeting).** *During the weeks before the final exam of a lecture, Sally regularly meets with two fellow students in an open space at the library. Together they review the lecture handout and their annotations and discuss unclear topics. If necessary, they collect further information from textbooks and web pages and link these to their lecture handouts. At the end of each meeting, they collaboratively create a summary of the topics. As usual, the other group members dictate and Sally writes the summary because she has the best handwriting of all members.*

First, users can collaborate in a *co-located* setting, for example in meetings. Several persons use the system at the same time in the same place. They interact with printed documents on paper and with digital documents on one or more pen-enabled displays (Fig. 5.4). Each user has his or her own pen. In this setting, CoScribe is designed for a rather small number of users (about two up to six users). In a small group, users can physically share their documents or work concurrently on the same document.

**Scenario 3 (Accessing Documents at Home).** *Dan is ill and could not attend today's learning group meeting. He uses the CoScribe viewer to access the summary document created by the other team members.*

Second, users can collaborate over the *distance*, for example to continue collaboration after a co-located meeting. They can then access digital versions of the documents (and print these, if desired). The documents include annotations, links and tags that were made by other users. In a remote setting, CoScribe supports a very large number of users, ranging up to several hundreds of users in a large lecture. While it provides some support for remote *synchronous* settings, *asynchronous* forms of remote collaboration are more adequate for user interfaces that heavily rely on



Figure 5.4: Co-located collaboration

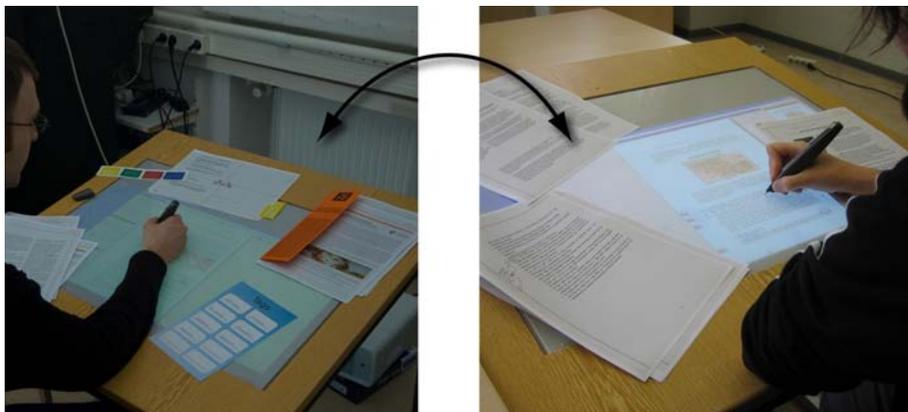


Figure 5.5: Remote collaboration

paper. This is due to the fact that contents on paper cannot be updated in real-time by the digital system. In contrast, it is not problematic to re-print updated versions once in a while.

In order to provide efficient remote access to shared contents of many users, several *collaborative visualizations* for documents, annotations, hyperlinks and tags have been developed which integrate user-generated content of several users.

- First, the user can print an updated version of the document which includes the annotations, links and tags of one or more other users.
- Second, user-generated content of other users is automatically integrated into the digital representation of a document, which is displayed in a software viewer.
- Third, an ecology-centered visualization integrates all documents and user-generated content from all users. It provides overview on and structured access to the entire collection of documents. This is particularly supportive if a considerable number of users collaborate on a large number of documents, a setting where it might be difficult to find a given document or a given passage of a document later on.
- Finally, further more specialized visualizations focus on specific aspects. For example, one view integrates the bookmarks made by all users on a particular document in order to support users in comparing the structuring created individually by each user.

## 5.2 Integrating Printed and Digital Documents

In this section, we take a closer look on the means by which CoScribe integrates the interaction with printed and with digital documents. Here, we discuss aspects that are of general importance for the entire system. In contrast, the subsequent sections will then elaborate on aspects which are specific to a particular activity (annotating, linking or tagging). Table 5.1 provides an overview on the challenges which we address in this chapter.

### 5.2.1 A Unified Interaction Device

**Traditional practice.** In traditional practice, knowledge workers typically dispose of both physical and digital tools for interacting with documents. This attributes to the fact that both printed and digital documents are used. The main tools for interacting with physical documents are paper and one or several pens (ballpoint pen, pencil, differently colored highlighter etc.) complemented by helper tools like erasers and rulers. Further physical tools that are frequently used support the organization and structuring of documents. This includes folders and ring binders, storage boxes, staplers, retaining clips, bookmarks and page marker stickers. Main tools

| Desideratum  | Approach  | Section | Concept | Function | Innovation w.r.t. related work | Main requirements |
|--|---|---------|---------|----------|--------------------------------|-------------------|
| One unified interaction device for paper and screens | Anoto-enabled screen  | 5.2.1   | •       | •        | •                              | R1a, R4a          |
| Mobile use and real-time interaction                 | Batch processing and streaming of pen data  | 5.2.2   | •       | •        |                                | R4                |
| Tangible interaction                                 | Printed documents and printed tools   | 5.2.3   | •       | •        | •                              | R1                |
| Digital integration                                  | Digital document viewer and plug-in for Mozilla Firefox   | 5.2.4   | •       | •        | •                              | R1                |
| Structured access to all information of the ecology  | Ecological view   | 5.2.5   | •       | •        | •                              | R1, R3, R4c       |
| Choice between printed and digital representations   | Printed buttons for accessing viewer / Printing documents; map-based or property-based indication | 5.2.6   | •       | •        |                                | R1                |

Table 5.1: Desiderata and our approaches concerning the integration of printed and digital documents



Figure 5.6: Using the same digital pen to interact on real paper and on displays

for interacting with digital documents are displays, keyboards and mice. There exist more specific devices like Tablet PCs, which imitate the use of pens on paper documents, and E-Book readers. However, these are far less widespread.

As discussed in Chapter 3, approaches proposed in the related work require the user to switch between different input devices for interacting with physical and digital documents. Typically, the input device for interacting with printed paper is a digital pen, whereas interaction with digital media takes place using keyboards and mice or pen-based displays with styluses other than the digital pen used on paper. Hence, we encounter not only a discontinuity between physical and digital media, but also a *discontinuity of tools*, as a given tool applies either to printed or to digital media. A small number of more advanced systems have been presented which use the same stylus for paper and digital media (e.g. [Wel93, EMBW08]). However, these systems require a complex technical setup including vision-based tracking of paper documents, which substantially restricts the mobility of paper.

**The same pen on paper and on displays.** CoScribe supports using the same pen both on printed media and on displays while maintaining the mobility of paper (Fig. 5.6, left and center). Almost all interactions can be made in a similar manner on printed documents or on digital documents. Some techniques even use pen gestures that span printed and digital documents, e.g. when creating a hyperlink between a printed document and a Web page. These interactions are mainly inspired from traditional practices of working with documents, but are complemented with elements from digital user interfaces.

On paper, the Anoto pen acts as a traditional ballpoint pen, which additionally performs a digital capturing of the ink traces. On the display, the digital pen behaves similarly to styluses of common pen-enabled displays, such as those of Tablet PCs, with the difference that several pens can be simultaneously used on the same display. For this purpose, we developed a specific display which supports input with one or several digital pens. It can be used as a tabletop display or as a vertical

screen. Section 6.2 describes this prototype in more detail. In addition to interacting with CoScribe, third-party applications can be controlled with the pen, as pen input on third-party applications is translated to ordinary mouse events.

### 5.2.2 Mobile Use and Real-Time Interaction

To support both autonomous mobile use and real-time interaction with the digital system, CoScribe offers two ways of synchronizing pen data with the digital system:

1. *Batch synchronization*: In order to maintain the mobility which is inherent to paper, the digital pen can be used autonomously. This means, that it does not require a nearby device, which for example supports tracking the location of the pen or receives data from the pen. During writing on paper, data is buffered on the pen. At regular intervals, when a computer is available, all data temporarily buffered on the pen is transferred to the digital system. Batch synchronization supports particularly well writing tasks, which do not require immediate system feedback. However, no real-time interaction with the digital system is possible.
2. *Real-time streaming*: If a computer is nearby, pen data can be sent in real-time to the digital system over a wireless Bluetooth channel. This has the advantage that all data can be immediately processed, which for example is important for providing real-time feedback when using the pen as a stylus on a display. However, the mobility of paper is restricted to an area within the range of the wireless connection. With the current Bluetooth specification, this is up to distances of about 100 metres.

### 5.2.3 Printed Documents and Tools

**Documents.** The CoScribe print tool enables the user to produce printed versions of *documents*. The tool adds the Anoto pattern to all printouts, which is required by Anoto pens for tracking pen traces. Although printed documents can be used like traditional documents, they have an underlying digital model. The system automatically adds changes made on printed documents to this digital model.

**Printed tools.** In addition to printed documents, CoScribe comprises various *tangible tools*, each serving specific purposes. Such tools are for example bookmarking stickers and leaf binders. In contrast to the tangible tools used in typical Tangible User Interfaces, our tools are made of paper and moreover do not require additional electronic components. For these reasons, an ordinary printer suffices to create a large number of tangible tools quickly and at very low cost. Figure 5.7 depicts several types of tangible tools which are used in CoScribe and that are presented in the following chapters.



Figure 5.7: Tangible tools which are made of paper

**Tools as objects of interest.** These tangible tools have the characteristic that they are not only instruments but also objects of interest. Even if they are not used as a tool, they have additional value for the user. We explain this in the following.

On the one hand, there are tools that have a purely instrumental function. This means that a tool does not have another purpose than helping the user in interacting with the system in order to accomplish a specific task. The tool itself is not an object of interest but is used to manipulate objects of interest (e.g. documents). Examples of traditional instrumental tools are hammers, scissors and pens. Typical instruments of Graphical User Interfaces, such as instruments for changing the font size, for drawing graphical elements and for scrolling within documents, do not have value besides their instrumental function.

On the other hand, in addition to this instrumental function, a tool can be an object of interest. This means that it has an own value even if it is not used to manipulate other objects of interest. We aim at designing tools that incorporate both functions. For example, an adhesive sticker that can be attached to a document page in order to bookmark this page is initially an instrument for creating an index on a document page. Once the sticker is attached to a document page, it becomes an object of interest, as it represents the actual index. Tag Menu Cards are instruments for defining tags and applying them to documents. Yet, the Tag Menu Card is also an inventory for looking up all previously defined tags of one user and therefore has an additional value for the user. Similarly, physical folders are used as an instrument to define collections of documents. In addition, the physical state of the folder indicates for example if only a few or many documents are included in this collection, the user can open the front flap and check up which documents are contained within it, etc. Therefore, it is itself an object of interest.

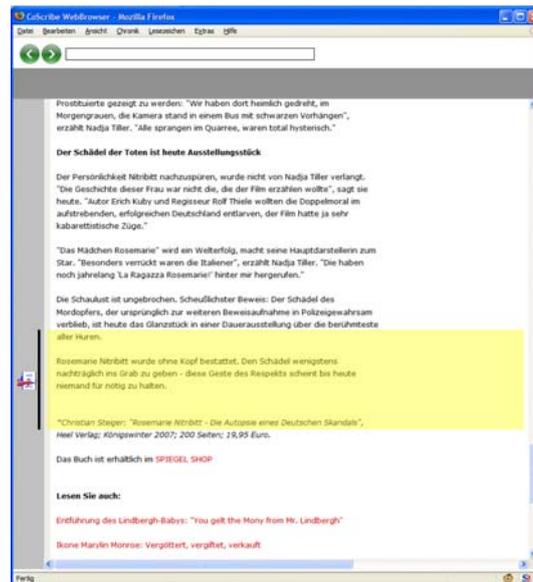


Figure 5.8: A CoScribe Firefox plug-in enables pen-based input on web pages and displays additional information (here information on hyperlinks)

### 5.2.4 Digital Document Viewer

While paper provides for a flexible and intuitive use of documents, there might be situations in which the user prefers working with a digital version of a document. For example, full text search, audio and video are better supported by digital documents. Moreover, digital visualizations enable to access shared comments of other users in real-time. For these reasons, CoScribe does not only support printed documents. The user can access a digital version of any document in a software viewer. This version includes own and shared annotations, hyperlinks and tags.

CoScribe includes several viewers for different types of documents. Web pages are available in Mozilla Firefox. A CoScribe Firefox plug-in enables pen-based input on web pages and displays additional information, such as hyperlinks created by the user (Fig. 5.8). For PDF and PowerPoint documents, CoScribe includes an own document viewer (Fig. 5.3 on p. 115). While we opted for a proprietary solution for our prototype, future versions of CoScribe could use plug-ins for standard software, such as Adobe Acrobat and Microsoft PowerPoint.

The viewers can be used on standard computers. However, if used on the Anoto-enabled tabletop display, they can be controlled with the same pen which is used on paper. The same pen gestures as on paper create and modify annotations, links and tags directly on the digital documents.

In contrast to changes made on the printed representation which are automatically transferred to the digital representation, changes made within the viewer are not automatically reflected on paper. If a corresponding printed representation is

needed, one can use the printout module to print an updated version of the document.

### 5.2.5 Ecological View

The document viewers display a visualization that corresponds to an *individual document*. However, as shown in Section 2.2, one of the main advantages of paper is that one can lay out multiple sheets and documents on the desk. This provides for a better overview on documents and their relations. We designed a digital visualization that corresponds to an augmented view on the *entire desk*. This so-called ecological view integrates all documents into one single interactive visualization. Its focus is on the relations between documents and the activities that users performed on them.

This view is basically a digital equivalent to the arrangement of multiple documents in the physical workspace. However, in contrast to related work (e.g. [Wel93]), it does not aim at simply mirroring the *physical* arrangement. Instead, it details on the *logical* relations between documents. Documents are not visualized near each other because their physical locations are close but because they are connected by a hyperlink, tagged with the same tag concept or used by the same user. This provides for a structured access to the information of document collections.

**Scenario 4** (Structured Access to Information). *During the last few weeks, Sally has frequently met her learning partners. Together, they have revised the lecture handouts, have read a lot of further documents and have written some new documents. Now that the final exam is arriving, Sally wants to review all information that is related to Molières comedies, which is one of the topics of the exam. She selects the tag “Molière” and has direct access to all information that was tagged with this label. Moreover, she seeks a document of which she does not remember the contents. But she knows that Dan has extensively worked on it during one of their first meetings. She therefore selects all documents which Dan has worked on and restricts the time period. Of the three remaining documents, she quickly finds the one she was looking for.*

### Elements of the Visualization

Due to the focus on relations, we opted for an interactive graph visualization. A screenshot of the view is given in Fig. 5.9 and will be explained in the following. Documents and document pages are represented by thumbnail images. Folders, tags and users are represented by icons. These nodes are connected by edges that visualize links, tags and user activities. We now provide more details on these elements:

- *Hybrid documents* (Fig. 5.9 (1)) are displayed with a thumbnail image of their first page and have a shadow that corresponds to their number of pages. For a

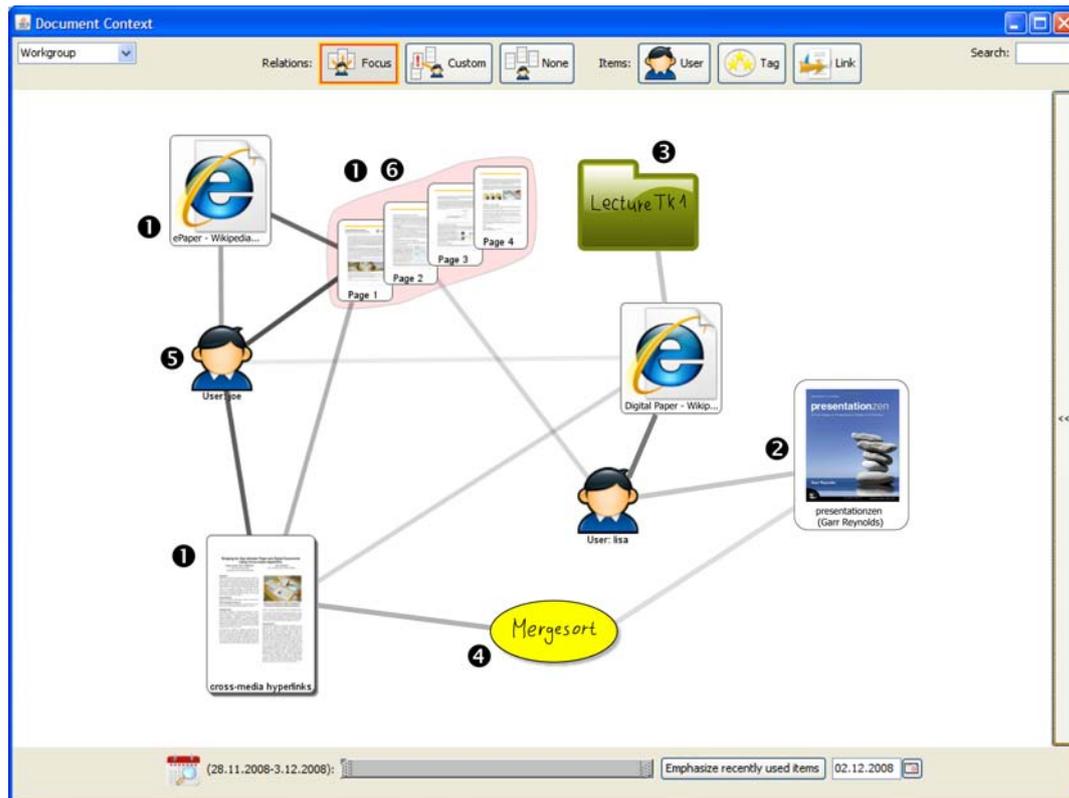


Figure 5.9: The ecological view

more detailed view, users can expand a document to thumbnails of its individual pages. Double clicking/two pen taps on a document or page thumbnail opens it the corresponding document viewer.

- *Printed-only documents* (2) are visualized by a thumbnail of the cover.
- *Folders* (3) are indicated with an icon in their color containing a handwritten label. A shadow corresponds to the number of pages of all documents contained in the folder. Like hybrid documents, folders can be “unfolded” to view the individual documents.
- *Tags* (4) are displayed as oval concept cards covered with the tag label.
- Each *user* (5) is represented by a personal icon.

An edge expresses:

- A *link* between two documents or folders. When clicking on the edge, the scope of the link is highlighted in yellow in the document thumbnail(s).

- *Tags* of documents or folders are indicated by an edge between the tag concept card and the document or folder.
- *User activities* are expressed by an edge from a user icon to all documents this user has annotated, linked, or tagged.

This graph provides a quick overview for example on how the collection of documents is structured by hyperlinks, on which user has worked on which documents or on which documents are tagged with a specific tag concept.

### Navigation and Filtering

A collaborative graph which contains the documents and activities of several users can quickly become complex. This makes high demands both on navigation and on appropriate reduction of the complexity. For this reason, the interactive visualizations enables users to navigate through a large graph by panning as well as to focus on various levels of details by zooming in and out and by collapsing or expanding documents and folders. In order to focus on an individual element, this can be automatically centered or moved by dragging it with the mouse or with the pen.

Moreover, the ecological view offers comprehensive filter options on five dimensions that can be intuitively manipulated:

- *Temporal filter*: A dual-end time slider (Fig. 5.10 (1)) enables to hide all items that have not been created or modified in a given period. To further allow a quick temporal overview of all items, a time mode varies the size of the nodes. Larger nodes represent items that have been created or modified recently while smaller nodes correspond to older items.
- *Global filter*: Hides or displays all links, all tags or all users. These filters are activated by modal buttons (Fig. 5.10 (2)).
- *Filter an individual user or an individual tag concept*: All documents which are modified only by this user (or tagged with only this concept) are hidden. This filter is activated by dragging the corresponding node and dropping it on a list of filtered elements which is displayed in a filter panel beneath the graph (Fig. 5.10 (3)). The user or the tag is then added to this list. By tapping or clicking on it, the documents are displayed again. It is important to note that not only one user or one tag can be filtered, but an arbitrary number of them.
- *Focus on one node*: Only this node and all directly adjacent nodes are displayed. One or more hidden nodes can be re-displayed by tapping or clicking on one or more entries in the filter panel.
- *Full-text search*: An arbitrary query string can be entered in a text input field (Fig. 5.10 (4)). All users, tags and documents that do not contain this string are hidden from the view.

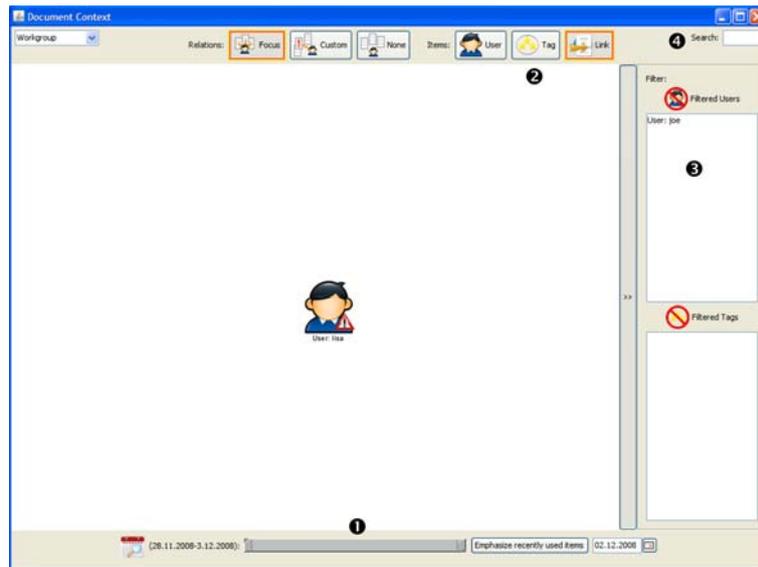


Figure 5.10: Filter options of the ecological view

### 5.2.6 Transformations between Printed and Digital Representations

A final issue, which concerns the entire system, is the question of how printouts and their corresponding digital views are coupled.

**Scenario 5** (Interwoven Use of Printed and Digital Representations). *After a lecture session, Sally usually briefly reviews the contents when she is back at home. She goes through the printed lecture scripts which includes the annotations she has made during the lecture. As she has difficulties understanding one particular slide, she wants to read the annotations made by the instructor and by her student fellows. For this purpose, she displays the digital version of this slide in the CoScribe viewer. This includes the annotations of other persons.*

*Then she browses in the digital version and reads other slides, as she hopes finding further helpful annotations of other students. After some time, she comes across a very helpful annotation on another slide. She decides to add this annotation directly to her printed script. She therefore searches the sheet of paper containing the slide and writes the annotation on it.*

Users can quickly access a digital version of a printed page by tapping with the pen on a button which is printed on each page (Fig. 5.11 upper right). The ecological view can be accessed with a second button (Fig. 5.11 upper left). The thumbnail of this particular document page is then centered and highlighted.

For the reverse direction – accessing a printed representation –, the user can print a new copy of the piece of information. However, it is impractical, expensive and harmful to the environment to re-print the information at a frequent basis. Another way of transformation is to access an already existing printed representation of this



tionelle Post-Its zur Indizierung von Papier-  
sind, können Interaktionen mit einem digi-  
tische Klassifikation mit vordefinierten Ty-  
ierten Titel. Alle Daten sind in Echtzeit im

Figure 5.11: These buttons for accessing the page in the document viewer (right) and in the ecological view (left) are printed in the top right corner of each document page.

piece of information. This implies that the user must find this physical copy, which can be hard if the user disposes of a large number of printed representations. Therefore, the digital system should provide support for finding an equivalent physical representation.

CoScribe provides support on two levels:

**Map-based indication.** In a non-mobile setting, CoScribe can digitally track the position of printed documents using ARTag [Fia04] marker-based location tracking. The location of a given document or document page is then highlighted on a photo of the desk.

**Property-based indication.** A drawback of map-based indication is that the technical setup for tracking the location of physical object strongly restricts mobile use. A second approach does not indicate the concrete physical location but provides information about properties of the representation, such as its visual appearance. The user can then utilize this information to find the physical representation. CoScribe uses the following properties:

- The *visual appearance* of the document (the thumbnail or document page view visually corresponds to the printed page). This is particularly helpful if the layout of different documents and pages is diverse. In this case, the size of the margins, the number of columns and the repartition of headings, paragraphs and illustrations can provide a good visual cue for finding a given document or page.
- If the document is contained within a *folder* (see p. 157), the system indicates this folder. As the number of folders is typically much smaller than the number of documents, this simplifies finding the document page.
- Digital Paper Bookmarks are adhesive sticker that the user can attach to individual document pages in order to index them. These stickers are well visible

because they juggle out of the document. The *arrangement of bookmarks* on a document as well as their colors and labels provide good indicators for identifying a specific document at a glance. More details on Digital Paper Bookmarks and the corresponding digital representations are given in Section 5.5.1.

Up to now, we have discussed issues relevant to the entire system that equally concern all the activities of annotating, linking and tagging documents. In the next three sections, we discuss interaction techniques and visualizations that are more specific for one of these three activities.

### 5.3 Annotating Documents

The interaction techniques and visualizations presented in this chapter aim at supporting one or several users to make handwritten annotations on printed and digital documents, to access these in both printed and digital form and to share annotations with co-workers.

On page 2, we defined an annotation as an amendment to an existing document that is conceptually separate yet contextually related. We define a specific sub-type, handwritten annotations, as follows:

**Definition 12** (Handwritten Annotations). *A handwritten annotation is an annotation that consists of a set of ink strokes which are spatially and temporally contiguous and made by one single user.*

The contents of handwritten annotations consist for example of an underlining, one or several words, formulae, drawings or combinations of these.

In the remainder of this chapter, we will first discuss the general interaction concept. This includes how annotations are made on which media and how information can be transformed between printed and digital representations. We will then briefly discuss how the user interface for annotating documents can be adapted to the user. In a next step, we focus on collaborative issues, particularly on a paper-based mechanism for sharing annotations and on novel visualizations of shared annotations made by other users. Finally, we will assess the performance of handwriting recognition and discuss to what extent handwriting recognition can support the interaction and visualization techniques. Table 5.2 provides an overview on the challenges which we address in this chapter.

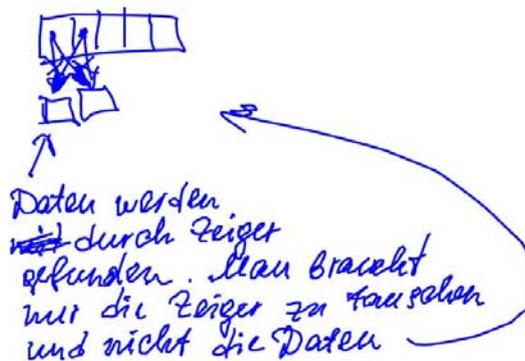


Figure 5.12: Example of an annotation which includes sketches

| Desideratum   | Approach  | Section | Concept | Function | Innovation w.r.t. related work | Main requirements |
|---|---|---------|---------|----------|--------------------------------|-------------------|
| Flexible free-form annotations                        | Unrestricted PPUi; printing documents with annotations                                      | 5.3.1   | •       | •        |                                | R1, R2, R4a       |
| User-adaptable PPUi                                   | Various print layouts   | 5.3.2   | •       | •        |                                | R4                |
|   | Free arrangements of several pages  | 5.3.2   | •       | •        | •                              | R4b               |
| Paper-based sharing and classification of annotations | Tagging with Buttons  | 5.3.3   | •       | •        | •                              | R1, R3, R4a       |
| Integrated access to shared annotations               | Integrated visualization using dynamic shrinking/expanding and repositioning of annotations | 5.3.4   | •       |          | •                              | R1, R3, R4a       |
|   | Integrated multi-user visualization   |         |         | •        | •                              |                   |
|   | Integrated multi-user printouts   |         |         | •        | •                              |                   |
| Use for presentations                                 | Presentation interface  | 5.3.5   | •       | •        |                                | R1, R3, R4        |
| Access to text of handwritten annotations             | Evaluation and use of handwriting recognition   | 5.3.6   | •       | •        |                                | R1, R3, R4        |

Table 5.2: Desiderata and our approaches concerning the annotation of printed and digital documents



Figure 5.13: Annotating a printed document

### 5.3.1 Handwritten Free-Form Annotations

**Scenario 6** (Mobile annotation). *Before the lecture, Sally prints the lecture script, which is provided online by the instructor. During the lecture, she then makes annotations on this script using her digital pen. She does not need a computer, but only the digital pen and the printout. As she lives at some distance from the university, she spends a considerable amount of time in public transport. On the way home, she typically uses this time to review the contents presented this day. Using the digital pen, she marks passages of the lecture script which are particularly important or which she has difficulties understanding.*

**Scenario 7** (Active reading). *Sally has the task to read an article until the next lecture. It is a challenging text, having a complex structure and argumentation and including many foreign words. Sally therefore works intensely on the text. During reading, she makes annotations: She underlines key words and marks up important passages. She looks up the signification of foreign words and annotates them with the corresponding English terms. Moreover, she adds clarifying details and notes her own thoughts and ideas concerning the text.*

Our design was guided by the goal of finding an appropriate trade-off between an interaction strategy which is as unconstrained as possible, permitting each user to maintain his or her personal annotation style (which is highly personal [Mar97]), and imposing some minimal formalization, which can be leveraged for computer interpretation of informal annotations. Moreover, we judge it important to maintain most of the affordances of pen and paper, such as its highly mobile use and the combined use of multiple sheets of paper. Therefore, the aim is to provide for a smooth transition from traditional paper annotations to a computer-supported collaborative annotation process. In order to further reduce the gap between paper and the digital world, the same interactions as on paper shall be possible on screens.

Users can print digital documents on paper and annotate them with the digital pen (Fig. 5.13). The printed user interface is designed to constrain the personal an-



Figure 5.14: A document containing printed annotations

notation style as little as possible. Handwritten annotations, including free-form content such as sketches or formulae, can be made anywhere on the printouts. At any time, the user can add, modify or delete annotations. An annotation is deleted by performing a cross-out gesture on the annotation. It is then removed in the digital viewer and not printed in subsequent printouts.

A print module allows printing documents on paper. In addition to the printed lecture slides, the pages can optionally contain free areas. Our observations have shown that free areas on lecture handouts seemed necessary to most students, because first the document pages often do not contain enough empty space for extensive annotations and second it is considered important to take notes independently of the instructor's structuring.

Moreover, a document can optionally include annotations that were made beforehand (Fig.5.14). This comprises own annotations, shared ones of another user or annotations originating from several users (more details will be given below). Further, the module can add visual information about annotations which was not visible on the printed instance where the original annotation was made. This includes the category of an annotation (color coding), the time when an annotation was created (older annotations are printed lighter) and the number of times an annotation was read by other users (annotation turning more red successively).

As analyzed in Section 3.3.2, this improves over previous work for paper-based annotation, which to the best of our knowledge is restricted to one fix layout, prints documents only with own annotations (if at all) and does not visualize additional

information.

The equipment which is needed for the interaction technique is highly mobile, lightweight and comparable to that used traditionally for notetaking. The user only requires paper printouts of the document(s) and a digital pen. Hence, the user is not bound to a specific location and the technology remains in the background as much as possible. Due to the Anoto technology used in CoScribe, it is possible to simultaneously work with several sheets of paper and to use the same pen on all of them without additional effort (like for example manually calibrating each sheet of paper or pressing on a button when switching between pages etc.). This provides for a very natural working style.

**Scenario 8** (Reviewing Annotations). *At home, she puts the pen into a reading device, which automatically transfers all pen data to the CoScribe client running on Sally's computer. She then consults a textbook and completes the notes she has made during the course. The pen is now used in live-mode, i.e. all data is available in CoScribe in real-time. As she has not well understood one particular slide, she opens the CoScribe document viewer and reads the annotations that other students have made on this slide.*

In addition to the use on paper, the same digital pen can be used to make annotations on digital documents which are displayed in the CoScribe viewer on the pen-enabled display (Fig. 5.6). As on paper, handwritten annotations can be made anywhere on the document. From a user's point of view, this is similar to annotating documents with a stylus on a Tablet PC, but one does not have to switch between different pens for paper and displays.

In contrast to changes made on the printed representation, which are automatically transferred to the digital representation, changes made within the viewer are not automatically reflected on paper. If a corresponding printed representation is needed, one can use the printout module to print an updated version of the document.

### 5.3.2 Adaptivity of the Printed User Interface

Graphical User Interfaces can be flexibly and immediately adapted to the user's needs that depend on the task to perform, the current context and his or her personal preferences. For example, both the user and the system can easily change the size of windows and their layout and can easily modify the position and the contents of individual interface elements, such as toolbar palettes. In contrast, paper is a more static medium. Once printed, the layout of a printed page is fix and moreover the digital system cannot easily update the contents.

Yet, qualitative findings of the evaluation of CoScribe (see Section 7.2) show that also in a paper environment, users prefer different layouts. For instance, right-handers prefer empty notetaking areas being located to the right of the actual document page, while left-handers prefer them to the left. While in some contexts, the user makes many annotations on the document (e.g. during a lecture), other contexts (e.g. excerpting on a separate sheet of paper) do not require making many

annotations and therefore, the pages of the document can be printed in a down-scaled manner. This challenges the design of Pen-and-Paper User Interfaces.

To address this issue, CoScribe offers a flexible printed user interface. The user can adapt the Pen-and-Paper User Interface of CoScribe to fit his or her personal preferences and the current context. This is supported both at the levels of individual pages and of the ensemble of pages used.

**Individual pages.** Users can choose among various layouts of printed document pages. For instance, each printed page can contain one page of the document or alternatively, several document pages can be printed on one single sheet of paper in a scaled-down layout. Optionally, if the document pages do not contain much space for making annotations, empty notetaking areas can be included. The position of other interface elements, such as button toolbars, can be modified as well. Figure 5.15 gives several example layouts. Moreover, documents can be re-printed in a different layout with all annotations correctly resized and repositioned. For this reason, our model stores annotations independently from the layout of the document. Similarly to printed document pages, the size and layout of other print products (such as bookmarking stickers and Tag Menu Cards) can be adapted by the user.

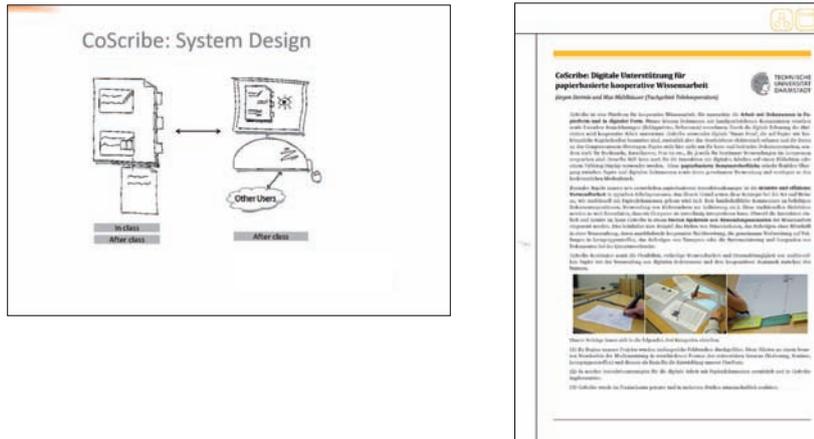
**Several pages.** Adapting the print layout implicates the drawback that the user has to make a decision before printing the document. Adapting the layout during use is hard, since this requires re-printing the document. For this reason, CoScribe offers a more dynamic way of modifying the printed user interface. In addition to adapting the layout of single pages before printing them, users can dynamically modify the layout built of several printed pages. If the space available on a document page (and on the optional notetaking area) is not sufficient for making annotations, users can dynamically add one or more further empty paper sheets to this document page by associating them with a line gesture (see Section 5.4.2 below). This corresponds to the GUI interactions of enlarging a window to have more space available or to scrolling in a document view to display an empty area.

### 5.3.3 Paper-based Sharing of Annotations

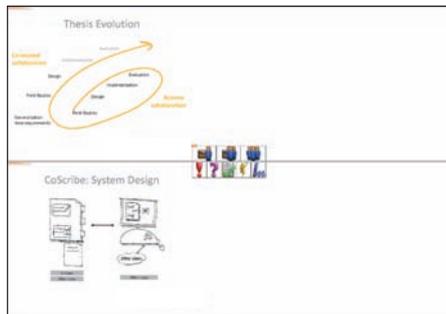
As discussed in the previous chapters, collaboration is an important element of document-based knowledge work. While paper documents are well-suited for co-located collaboration in many aspects, they constrain remote collaboration (in comparison to digital documents) because their contents are only available physically and not digitally.

Section 4.4.1 above described how CoScribe can be used in co-located collaborative settings. Each user disposes of his or her personal pen. This enables to attribute the activities of individual users to this user. Moreover, several pens can be used in the same time on one single sheet of paper on one display. Further, the users can work on multiple documents at a time.

One page per sheet



Two pages per sheet



Four pages per sheet

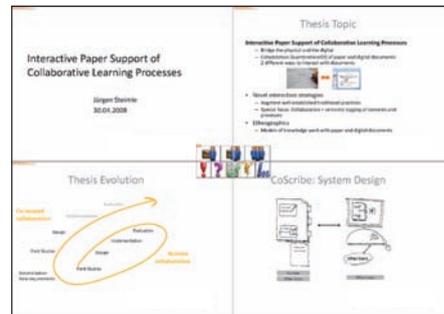


Figure 5.15: Example layouts of printouts

This section focuses on the second collaborative setting: remote sharing of annotations. CoScribe supports asynchronous collaboration over distance. It enables users to share their annotations with collaborators. These can access shared annotations in their document viewers and/or print the document including these annotations. This section discusses an unobtrusive interaction technique that allows the user to classify the visibility of annotations directly when writing the annotation on paper.

**Scenario 9** (Private and Public Annotations). *It is absolutely acceptable for Sally to share her annotations of the lecture script with the other members of her learning group. Sometimes she makes off-topic notes, which she prefers to remain personal. For example, she makes an appointment with a fellow student and notes his phone number or she notes what she wants to buy after the lecture. For this reason, she marks these notes as private. They are not shared with other persons.*

**Interaction technique.** CoScribe offers three levels of visibility with which the user can classify individual annotations:

- *Private visibility:* The content is not shared with other users.
- *Group visibility:* Each user can set up groups with several other users. In different contexts (e.g. a specific lecture, a seminar etc.), the user can be member of different groups. Content with group visibility is shared with the members of all groups of this user that apply to the given context.
- *Public visibility:* The content is visible to all users in this given context (e.g. all students attending this lecture, all participants of the seminar etc.).

Our goal was to provide an interaction technique which can be seamlessly integrated with annotating and which is quick and reliable. Related research [LGA<sup>+</sup>07, LGHH08] discusses several means for classifying annotations suggesting spatial differentiation, pen differentiation and differentiation with pen gestures. Providing on the printout separate areas for each visibility level and requiring the user to write a note in the corresponding area is impractical for annotations, which often have to be made at specific context positions on the slides. Using a different pen for each level is intuitive but requires extra hardware. Moreover, research shows that students rather use one single pen than switching among many marking tools [Mar97]. A third solution consists in classifying notes by performing specific pen gestures. However, current digital pens cannot recognize gestures by themselves. Gesture recognition would have to be performed by the back-end system. Hence, it is not possible to provide feedback on success or failure of the gesture recognition in the mobile setting where pen data is not immediately transferred to the back-end system. Moreover, the system would have to distinguish gestures from ordinary handwritings or drawings. For this purpose, the cited research papers suggested to use additional hardware like a foot pedal or a second pen. Novel generations of digital pens which include gesture recognition and buttons could solve these problems in the future.

We therefore propose a fourth concept, which is inspired by buttons in Graphical User Interfaces because the interaction of pressing button is quick, easy and reliable. For this purpose, a toolbar containing several printed “buttons” is printed in the center region of each paper sheet (see Fig. 5.16). Each button represents one visibility level. A visibility is associated to an individual annotation by tapping with the pen on the corresponding button before or after writing the annotation. Moreover, a visibility level can be set or modified later on by making two consecutive pen taps on the button and on the annotation. While no graphical feedback on the tagging is provided on the printed slide unless the user makes additional markings, the visibility level is visualized with specific colors in the CoScribe viewer and on subsequent printouts. The viewer contains similar buttons as the printouts for defining or modifying visibilities. The same interaction technique is used for tagging annotations with semantic categories.

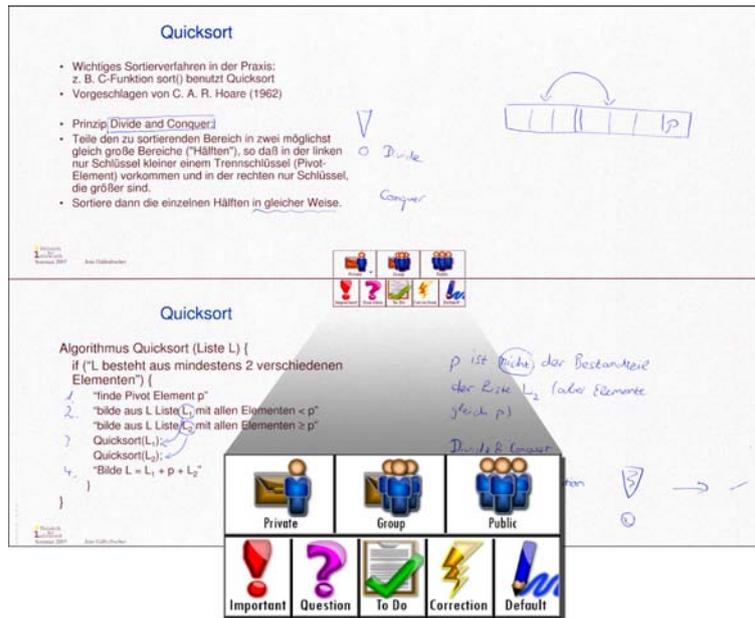


Figure 5.16: A button toolbar is printed on each page. These buttons provide for defining the visibility level of annotations (upper buttons) and for tagging them with semantic types (lower buttons)

**Discussion.** Performing a simple pen tap is quick and can be easily included into the annotation process. Moreover, defining visibility is optional, allowing the user to maintain a natural annotation style. If no visibility is chosen, the default level defined by the user is set. This reduces extraneous cognitive load during the annotation process. According to the interviews conducted during the evaluations, a very appropriate default level is the group level. With respect to privacy, it is typically not considered critical to share annotations with other members of the own group, as these are personally selected by the user. With this default level, only a small number of annotations that contain private information or that shall be visible to all users must be explicitly classified with a visibility level. If appropriate, the interaction technique could also support less or more than three levels or more than one group the user belongs to.

Applying visibilities to entire annotations instead of unstructured sets of pen strokes requires clustering pen strokes into annotations. For this reason, CoScribe uses a clustering algorithm for handwritten input that relies on temporal and spatial information. Possible errors can be manually corrected by the users. The software viewer therefore provides two functions for splitting annotations and merging pen strokes into one annotation.

In contrast to gesture-based differentiation, our approach is faster and reliable even with current pens that do not include displays nor processing power for ges-



Figure 5.17: A LED (within the red circle) provides feedback when the pen taps on a printed area

ture recognition. Moreover, it requires only one single pen. A correct interpretation is guaranteed, as determining the pen position on a paper button does not imply uncertainty. Current pens moreover provide graphical feedback to the user at the moment the button area is tapped on (Fig. 5.17). Hence, the user can be sure the classification has been correctly recognized.

A problem of current Anoto pens is that they cannot provide feedback on the current mode. For this reason, changes in the classification by selecting a button do only apply to the annotation created immediately afterwards, then the system returns to the default classification state. Thus the user does not have to remember a current system state over a longer period of time. If she desires absolute certainty to be in the default mode, she can tap on a 'default' button. While a completely modeless design would be preferable, this "semi-modal" design accounts for the two stages of selecting an instrument and an operand and nevertheless copes with the absence of feedback. Novel generations of digital pens that include a display<sup>1</sup> will solve this problem and make the use of button-based classification more comfortable.

#### 5.3.4 Visualization of Shared Annotations

This section describes how shared annotations of other users can be accessed with CoScribe. The following scenarios provide several examples how individual learning can be supported by shared annotations.

---

<sup>1</sup>A first pen including a display [Liv] is now commercially available. However, it does not support real-time streaming.

**Scenario 10** (Thoughts of other Persons). *Sally has not well understood a particular issue. Reading the annotations of other persons helps her in getting the idea. In another course, a controversial topic was discussed. She is interested in what the other students think of this point and reads their annotations.*

**Scenario 11** (Catching up on a Lecture). *Sally was ill for some days and missed several lectures. She therefore carefully reads the annotations that her learning partners have made during the lectures.*

**Scenario 12** (Collaborative Notetaking). *Sally attends a course in which the handout is not very detailed. For this reason, the students take very extensive notes and sometimes there is not sufficient time to note everything which is important. Hence, Sally and two other students have agreed to jointly take notes. Each person is responsible for one particular aspect and notes everything which is related to this aspect. After the course, each of them completes his or her own notes with the notes of the two other students. This leads to more comprehensive notes.*

In a traditional printed document, collaborative handwritten annotations are well visible because they are all made on the *same* physical document and the annotators can pay attention that their annotations do not overlap. Digitizing annotations enables the possibility to integrate annotations that are made on *different* physical copies of one document into one single digital version of the document. This stands in contrast to having a separate digital document for each physical copy. However, this poses the challenge of how the annotations of several users can be visualized in an aggregated view. As the users have different physical copies, their annotations might overlap because they are not aware of where other users make their annotations. An aggregated view would quickly become cluttered or even illegible when displaying a large number of possibly overlapping annotations of different users directly on the document. In the following, we discuss how we face this problem with digital and printed aggregated views.

**Digital representation.** The CoScribe viewer provides digital access to both own and shared annotations.

By separating the annotations of different users into different views, each of these views in itself becomes easier to read. The CoScribe viewer includes a *single-user view* for each member of the user's learning group. This view displays the annotations of this specific user. Yet, this implies the need to manually switch between views of different users. This becomes particularly cumbersome in larger communities.

In order to combine both one's own and shared annotations in an integrated integrated *multi-user view*, one has to deal with overlapping annotations. We examined three techniques for addressing this problem. First, overlapping annotations could be moved to the margin of the documents. This implies the drawback of separating annotations from their context. Annotations that visually refer to the context (e.g.

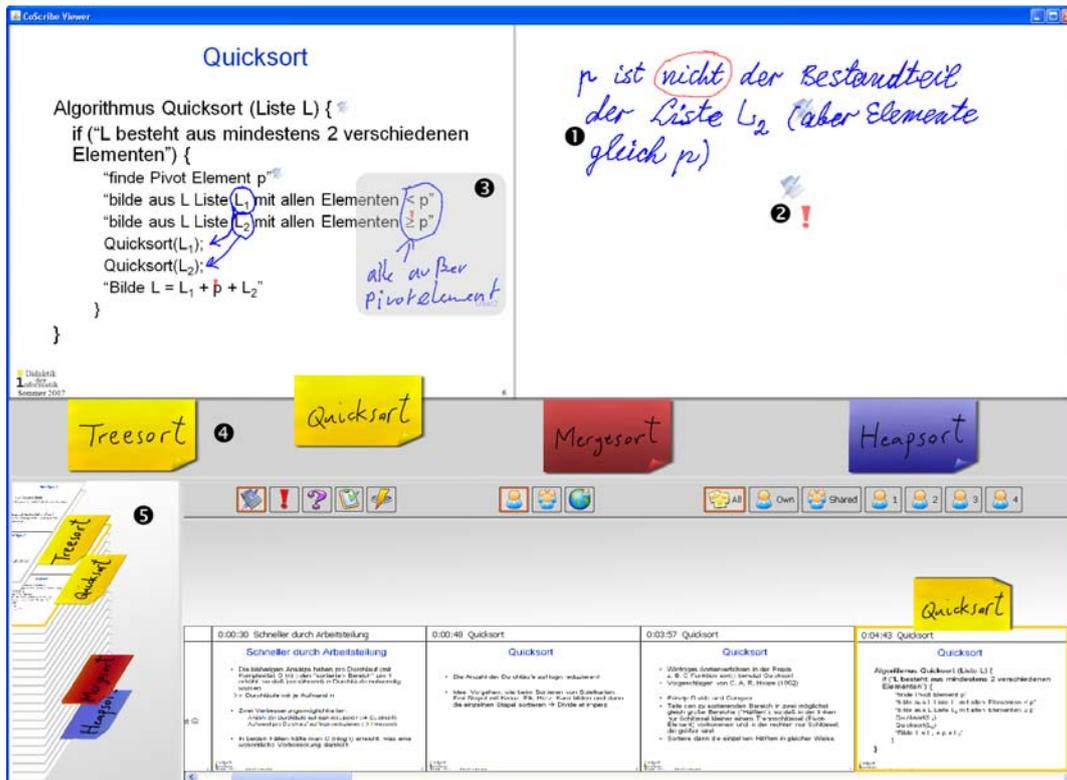


Figure 5.18: The CoScribe document viewer

underlinings, arrows, captions for printed elements) can become illegible. Second, the white spaces within a document page could be stretched to provide enough space for all annotations. However, this may result in very large document pages if the number of annotations is high and moreover overlapping annotations that refer to the same elements might lose their context, similarly to moving annotations to the margins.

CoScribe relies on a third technique. This consists in varying the size of individual annotations: collapsing annotations that are currently not relevant and expanding annotations that are in the user's focus. In the CoScribe viewer, one's own annotations are visualized as they are written on paper, whereas shared comments of other users are displayed in a condensed form. Instead of the annotation itself, a small icon is visualized at the position of the annotation (Fig. 5.18 (1)). This icon corresponds to the annotation category and varies in size according to the size of the annotation. When hovering with the mouse over the icon or tapping with the pen on it, the annotation is expanded and displayed at the correct position in its original size (Fig. 5.18 (2), annotation with grey background). The user can copy shared annotations considered especially relevant to his or her own script where they are displayed in their decompressed form like one's own annotations. More-

**Frankfurter Skandal der 50er**

**Der Mordfall der Rosemarie Nitribitt**

Nitribitt wurde mit einer Platzwunde am Kopf und Würgemark am Hals tot in ihrer Wohnung in Frankfurt am Main in der Stiftstr. 36 am Eschenheimer Turm, <sup>1957</sup> <sup>1958</sup> <sup>1959</sup> <sup>1960</sup> <sup>1961</sup> <sup>1962</sup> <sup>1963</sup> <sup>1964</sup> <sup>1965</sup> <sup>1966</sup> <sup>1967</sup> <sup>1968</sup> <sup>1969</sup> <sup>1970</sup> <sup>1971</sup> <sup>1972</sup> <sup>1973</sup> <sup>1974</sup> <sup>1975</sup> <sup>1976</sup> <sup>1977</sup> <sup>1978</sup> <sup>1979</sup> <sup>1980</sup> <sup>1981</sup> <sup>1982</sup> <sup>1983</sup> <sup>1984</sup> <sup>1985</sup> <sup>1986</sup> <sup>1987</sup> <sup>1988</sup> <sup>1989</sup> <sup>1990</sup> <sup>1991</sup> <sup>1992</sup> <sup>1993</sup> <sup>1994</sup> <sup>1995</sup> <sup>1996</sup> <sup>1997</sup> <sup>1998</sup> <sup>1999</sup> <sup>2000</sup> <sup>2001</sup> <sup>2002</sup> <sup>2003</sup> <sup>2004</sup> <sup>2005</sup> <sup>2006</sup> <sup>2007</sup> <sup>2008</sup> <sup>2009</sup> <sup>2010</sup> <sup>2011</sup> <sup>2012</sup> <sup>2013</sup> <sup>2014</sup> <sup>2015</sup> <sup>2016</sup> <sup>2017</sup> <sup>2018</sup> <sup>2019</sup> <sup>2020</sup> <sup>2021</sup> <sup>2022</sup> <sup>2023</sup> <sup>2024</sup> <sup>2025</sup> <sup>2026</sup> <sup>2027</sup> <sup>2028</sup> <sup>2029</sup> <sup>2030</sup> <sup>2031</sup> <sup>2032</sup> <sup>2033</sup> <sup>2034</sup> <sup>2035</sup> <sup>2036</sup> <sup>2037</sup> <sup>2038</sup> <sup>2039</sup> <sup>2040</sup> <sup>2041</sup> <sup>2042</sup> <sup>2043</sup> <sup>2044</sup> <sup>2045</sup> <sup>2046</sup> <sup>2047</sup> <sup>2048</sup> <sup>2049</sup> <sup>2050</sup> <sup>2051</sup> <sup>2052</sup> <sup>2053</sup> <sup>2054</sup> <sup>2055</sup> <sup>2056</sup> <sup>2057</sup> <sup>2058</sup> <sup>2059</sup> <sup>2060</sup> <sup>2061</sup> <sup>2062</sup> <sup>2063</sup> <sup>2064</sup> <sup>2065</sup> <sup>2066</sup> <sup>2067</sup> <sup>2068</sup> <sup>2069</sup> <sup>2070</sup> <sup>2071</sup> <sup>2072</sup> <sup>2073</sup> <sup>2074</sup> <sup>2075</sup> <sup>2076</sup> <sup>2077</sup> <sup>2078</sup> <sup>2079</sup> <sup>2080</sup> <sup>2081</sup> <sup>2082</sup> <sup>2083</sup> <sup>2084</sup> <sup>2085</sup> <sup>2086</sup> <sup>2087</sup> <sup>2088</sup> <sup>2089</sup> <sup>2090</sup> <sup>2091</sup> <sup>2092</sup> <sup>2093</sup> <sup>2094</sup> <sup>2095</sup> <sup>2096</sup> <sup>2097</sup> <sup>2098</sup> <sup>2099</sup> <sup>2100</sup> <sup>2101</sup> <sup>2102</sup> <sup>2103</sup> <sup>2104</sup> <sup>2105</sup> <sup>2106</sup> <sup>2107</sup> <sup>2108</sup> <sup>2109</sup> <sup>2110</sup> <sup>2111</sup> <sup>2112</sup> <sup>2113</sup> <sup>2114</sup> <sup>2115</sup> <sup>2116</sup> <sup>2117</sup> <sup>2118</sup> <sup>2119</sup> <sup>2120</sup> <sup>2121</sup> <sup>2122</sup> <sup>2123</sup> <sup>2124</sup> <sup>2125</sup> <sup>2126</sup> <sup>2127</sup> <sup>2128</sup> <sup>2129</sup> <sup>2130</sup> <sup>2131</sup> <sup>2132</sup> <sup>2133</sup> <sup>2134</sup> <sup>2135</sup> <sup>2136</sup> <sup>2137</sup> <sup>2138</sup> <sup>2139</sup> <sup>2140</sup> <sup>2141</sup> <sup>2142</sup> <sup>2143</sup> <sup>2144</sup> <sup>2145</sup> <sup>2146</sup> <sup>2147</sup> <sup>2148</sup> <sup>2149</sup> <sup>2150</sup> <sup>2151</sup> <sup>2152</sup> <sup>2153</sup> <sup>2154</sup> <sup>2155</sup> <sup>2156</sup> <sup>2157</sup> <sup>2158</sup> <sup>2159</sup> <sup>2160</sup> <sup>2161</sup> <sup>2162</sup> <sup>2163</sup> <sup>2164</sup> <sup>2165</sup> <sup>2166</sup> <sup>2167</sup> <sup>2168</sup> <sup>2169</sup> <sup>2170</sup> <sup>2171</sup> <sup>2172</sup> <sup>2173</sup> <sup>2174</sup> <sup>2175</sup> <sup>2176</sup> <sup>2177</sup> <sup>2178</sup> <sup>2179</sup> <sup>2180</sup> <sup>2181</sup> <sup>2182</sup> <sup>2183</sup> <sup>2184</sup> <sup>2185</sup> <sup>2186</sup> <sup>2187</sup> <sup>2188</sup> <sup>2189</sup> <sup>2190</sup> <sup>2191</sup> <sup>2192</sup> <sup>2193</sup> <sup>2194</sup> <sup>2195</sup> <sup>2196</sup> <sup>2197</sup> <sup>2198</sup> <sup>2199</sup> <sup>2200</sup> <sup>2201</sup> <sup>2202</sup> <sup>2203</sup> <sup>2204</sup> <sup>2205</sup> <sup>2206</sup> <sup>2207</sup> <sup>2208</sup> <sup>2209</sup> <sup>2210</sup> <sup>2211</sup> <sup>2212</sup> <sup>2213</sup> <sup>2214</sup> <sup>2215</sup> <sup>2216</sup> <sup>2217</sup> <sup>2218</sup> <sup>2219</sup> <sup>2220</sup> <sup>2221</sup> <sup>2222</sup> <sup>2223</sup> <sup>2224</sup> <sup>2225</sup> <sup>2226</sup> <sup>2227</sup> <sup>2228</sup> <sup>2229</sup> <sup>2230</sup> <sup>2231</sup> <sup>2232</sup> <sup>2233</sup> <sup>2234</sup> <sup>2235</sup> <sup>2236</sup> <sup>2237</sup> <sup>2238</sup> <sup>2239</sup> <sup>2240</sup> <sup>2241</sup> <sup>2242</sup> <sup>2243</sup> <sup>2244</sup> <sup>2245</sup> <sup>2246</sup> <sup>2247</sup> <sup>2248</sup> <sup>2249</sup> <sup>2250</sup> <sup>2251</sup> <sup>2252</sup> <sup>2253</sup> <sup>2254</sup> <sup>2255</sup> <sup>2256</sup> <sup>2257</sup> <sup>2258</sup> <sup>2259</sup> <sup>2260</sup> <sup>2261</sup> <sup>2262</sup> <sup>2263</sup> <sup>2264</sup> <sup>2265</sup> <sup>2266</sup> <sup>2267</sup> <sup>2268</sup> <sup>2269</sup> <sup>2270</sup> <sup>2271</sup> <sup>2272</sup> <sup>2273</sup> <sup>2274</sup> <sup>2275</sup> <sup>2276</sup> <sup>2277</sup> <sup>2278</sup> <sup>2279</sup> <sup>2280</sup> <sup>2281</sup> <sup>2282</sup> <sup>2283</sup> <sup>2284</sup> <sup>2285</sup> <sup>2286</sup> <sup>2287</sup> <sup>2288</sup> <sup>2289</sup> <sup>2290</sup> <sup>2291</sup> <sup>2292</sup> <sup>2293</sup> <sup>2294</sup> <sup>2295</sup> <sup>2296</sup> <sup>2297</sup> <sup>2298</sup> <sup>2299</sup> <sup>2300</sup> <sup>2301</sup> <sup>2302</sup> <sup>2303</sup> <sup>2304</sup> <sup>2305</sup> <sup>2306</sup> <sup>2307</sup> <sup>2308</sup> <sup>2309</sup> <sup>2310</sup> <sup>2311</sup> <sup>2312</sup> <sup>2313</sup> <sup>2314</sup> <sup>2315</sup> <sup>2316</sup> <sup>2317</sup> <sup>2318</sup> <sup>2319</sup> <sup>2320</sup> <sup>2321</sup> <sup>2322</sup> <sup>2323</sup> <sup>2324</sup> <sup>2325</sup> <sup>2326</sup> <sup>2327</sup> <sup>2328</sup> <sup>2329</sup> <sup>2330</sup> <sup>2331</sup> <sup>2332</sup> <sup>2333</sup> <sup>2334</sup> <sup>2335</sup> <sup>2336</sup> <sup>2337</sup> <sup>2338</sup> <sup>2339</sup> <sup>2340</sup> <sup>2341</sup> <sup>2342</sup> <sup>2343</sup> <sup>2344</sup> <sup>2345</sup> <sup>2346</sup> <sup>2347</sup> <sup>2348</sup> <sup>2349</sup> <sup>2350</sup> <sup>2351</sup> <sup>2352</sup> <sup>2353</sup> <sup>2354</sup> <sup>2355</sup> <sup>2356</sup> <sup>2357</sup> <sup>2358</sup> <sup>2359</sup> <sup>2360</sup> <sup>2361</sup> <sup>2362</sup> <sup>2363</sup> <sup>2364</sup> <sup>2365</sup> <sup>2366</sup> <sup>2367</sup> <sup>2368</sup> <sup>2369</sup> <sup>2370</sup> <sup>2371</sup> <sup>2372</sup> <sup>2373</sup> <sup>2374</sup> <sup>2375</sup> <sup>2376</sup> <sup>2377</sup> <sup>2378</sup> <sup>2379</sup> <sup>2380</sup> <sup>2381</sup> <sup>2382</sup> <sup>2383</sup> <sup>2384</sup> <sup>2385</sup> <sup>2386</sup> <sup>2387</sup> <sup>2388</sup> <sup>2389</sup> <sup>2390</sup> <sup>2391</sup> <sup>2392</sup> <sup>2393</sup> <sup>2394</sup> <sup>2395</sup> <sup>2396</sup> <sup>2397</sup> <sup>2398</sup> <sup>2399</sup> <sup>2400</sup> <sup>2401</sup> <sup>2402</sup> <sup>2403</sup> <sup>2404</sup> <sup>2405</sup> <sup>2406</sup> <sup>2407</sup> <sup>2408</sup> <sup>2409</sup> <sup>2410</sup> <sup>2411</sup> <sup>2412</sup> <sup>2413</sup> <sup>2414</sup> <sup>2415</sup> <sup>2416</sup> <sup>2417</sup> <sup>2418</sup> <sup>2419</sup> <sup>2420</sup> <sup>2421</sup> <sup>2422</sup> <sup>2423</sup> <sup>2424</sup> <sup>2425</sup> <sup>2426</sup> <sup>2427</sup> <sup>2428</sup> <sup>2429</sup> <sup>2430</sup> <sup>2431</sup> <sup>2432</sup> <sup>2433</sup> <sup>2434</sup> <sup>2435</sup> <sup>2436</sup> <sup>2437</sup> <sup>2438</sup> <sup>2439</sup> <sup>2440</sup> <sup>2441</sup> <sup>2442</sup> <sup>2443</sup> <sup>2444</sup> <sup>2445</sup> <sup>2446</sup> <sup>2447</sup> <sup>2448</sup> <sup>2449</sup> <sup>2450</sup> <sup>2451</sup> <sup>2452</sup> <sup>2453</sup> <sup>2454</sup> <sup>2455</sup> <sup>2456</sup> <sup>2457</sup> <sup>2458</sup> <sup>2459</sup> <sup>2460</sup> <sup>2461</sup> <sup>2462</sup> <sup>2463</sup> <sup>2464</sup> <sup>2465</sup> <sup>2466</sup> <sup>2467</sup> <sup>2468</sup> <sup>2469</sup> <sup>2470</sup> <sup>2471</sup> <sup>2472</sup> <sup>2473</sup> <sup>2474</sup> <sup>2475</sup> <sup>2476</sup> <sup>2477</sup> <sup>2478</sup> <sup>2479</sup> <sup>2480</sup> <sup>2481</sup> <sup>2482</sup> <sup>2483</sup> <sup>2484</sup> <sup>2485</sup> <sup>2486</sup> <sup>2487</sup> <sup>2488</sup> <sup>2489</sup> <sup>2490</sup> <sup>2491</sup> <sup>2492</sup> <sup>2493</sup> <sup>2494</sup> <sup>2495</sup> <sup>2496</sup> <sup>2497</sup> <sup>2498</sup> <sup>2499</sup> <sup>2500</sup> <sup>2501</sup> <sup>2502</sup> <sup>2503</sup> <sup>2504</sup> <sup>2505</sup> <sup>2506</sup> <sup>2507</sup> <sup>2508</sup> <sup>2509</sup> <sup>2510</sup> <sup>2511</sup> <sup>2512</sup> <sup>2513</sup> <sup>2514</sup> <sup>2515</sup> <sup>2516</sup> <sup>2517</sup> <sup>2518</sup> <sup>2519</sup> <sup>2520</sup> <sup>2521</sup> <sup>2522</sup> <sup>2523</sup> <sup>2524</sup> <sup>2525</sup> <sup>2526</sup> <sup>2527</sup> <sup>2528</sup> <sup>2529</sup> <sup>2530</sup> <sup>2531</sup> <sup>2532</sup> <sup>2533</sup> <sup>2534</sup> <sup>2535</sup> <sup>2536</sup> <sup>2537</sup> <sup>2538</sup> <sup>2539</sup> <sup>2540</sup> <sup>2541</sup> <sup>2542</sup> <sup>2543</sup> <sup>2544</sup> <sup>2545</sup> <sup>2546</sup> <sup>2547</sup> <sup>2548</sup> <sup>2549</sup> <sup>2550</sup> <sup>2551</sup> <sup>2552</sup> <sup>2553</sup> <sup>2554</sup> <sup>2555</sup> <sup>2556</sup> <sup>2557</sup> <sup>2558</sup> <sup>2559</sup> <sup>2560</sup> <sup>2561</sup> <sup>2562</sup> <sup>2563</sup> <sup>2564</sup> <sup>2565</sup> <sup>2566</sup> <sup>2567</sup> <sup>2568</sup> <sup>2569</sup> <sup>2570</sup> <sup>2571</sup> <sup>2572</sup> <sup>2573</sup> <sup>2574</sup> <sup>2575</sup> <sup>2576</sup> <sup>2577</sup> <sup>2578</sup> <sup>2579</sup> <sup>2580</sup> <sup>2581</sup> <sup>2582</sup> <sup>2583</sup> <sup>2584</sup> <sup>2585</sup> <sup>2586</sup> <sup>2587</sup> <sup>2588</sup> <sup>2589</sup> <sup>2590</sup> <sup>2591</sup> <sup>2592</sup> <sup>2593</sup> <sup>2594</sup> <sup>2595</sup> <sup>2596</sup> <sup>2597</sup> <sup>2598</sup> <sup>2599</sup> <sup>2600</sup> <sup>2601</sup> <sup>2602</sup> <sup>2603</sup> <sup>2604</sup> <sup>2605</sup> <sup>2606</sup> <sup>2607</sup> <sup>2608</sup> <sup>2609</sup> <sup>2610</sup> <sup>2611</sup> <sup>2612</sup> <sup>2613</sup> <sup>2614</sup> <sup>2615</sup> <sup>2616</sup> <sup>2617</sup> <sup>2618</sup> <sup>2619</sup> <sup>2620</sup> <sup>2621</sup> <sup>2622</sup> <sup>2623</sup> <sup>2624</sup> <sup>2625</sup> <sup>2626</sup> <sup>2627</sup> <sup>2628</sup> <sup>2629</sup> <sup>2630</sup> <sup>2631</sup> <sup>2632</sup> <sup>2633</sup> <sup>2634</sup> <sup>2635</sup> <sup>2636</sup> <sup>2637</sup> <sup>2638</sup> <sup>2639</sup> <sup>2640</sup> <sup>2641</sup> <sup>2642</sup> <sup>2643</sup> <sup>2644</sup> <sup>2645</sup> <sup>2646</sup> <sup>2647</sup> <sup>2648</sup> <sup>2649</sup> <sup>2650</sup> <sup>2651</sup> <sup>2652</sup> <sup>2653</sup> <sup>2654</sup> <sup>2655</sup> <sup>2656</sup> <sup>2657</sup> <sup>2658</sup> <sup>2659</sup> <sup>2660</sup> <sup>2661</sup> <sup>2662</sup> <sup>2663</sup> <sup>2664</sup> <sup>2665</sup> <sup>2666</sup> <sup>2667</sup> <sup>2668</sup> <sup>2669</sup> <sup>2670</sup> <sup>2671</sup> <sup>2672</sup> <sup>2673</sup> <sup>2674</sup> <sup>2675</sup> <sup>2676</sup> <sup>2677</sup> <sup>2678</sup> <sup>2679</sup> <sup>2680</sup> <sup>2681</sup> <sup>2682</sup> <sup>2683</sup> <sup>2684</sup> <sup>2685</sup> <sup>2686</sup> <sup>2687</sup> <sup>2688</sup> <sup>2689</sup> <sup>2690</sup> <sup>2691</sup> <sup>2692</sup> <sup>2693</sup> <sup>2694</sup> <sup>2695</sup> <sup>2696</sup> <sup>2697</sup> <sup>2698</sup> <sup>2699</sup> <sup>2700</sup> <sup>2701</sup> <sup>2702</sup> <sup>2703</sup> <sup>2704</sup> <sup>2705</sup> <sup>2706</sup> <sup>2707</sup> <sup>2708</sup> <sup>2709</sup> <sup>2710</sup> <sup>2711</sup> <sup>2712</sup> <sup>2713</sup> <sup>2714</sup> <sup>2715</sup> <sup>2716</sup> <sup>2717</sup> <sup>2718</sup> <sup>2719</sup> <sup>2720</sup> <sup>2721</sup> <sup>2722</sup> <sup>2723</sup> <sup>2724</sup> <sup>2725</sup> <sup>2726</sup> <sup>2727</sup> <sup>2728</sup> <sup>2729</sup> <sup>2730</sup> <sup>2731</sup> <sup>2732</sup> <sup>2733</sup> <sup>2734</sup> <sup>2735</sup> <sup>2736</sup> <sup>2737</sup> <sup>2738</sup> <sup>2739</sup> <sup>2740</sup> <sup>2741</sup> <sup>2742</sup> <sup>2743</sup> <sup>2744</sup> <sup>2745</sup> <sup>2746</sup> <sup>2747</sup> <sup>2748</sup> <sup>2749</sup> <sup>2750</sup> <sup>2751</sup> <sup>2752</sup> <sup>2753</sup> <sup>2754</sup> <sup>2755</sup> <sup>2756</sup> <sup>2757</sup> <sup>2758</sup> <sup>2759</sup> <sup>2760</sup> <sup>2761</sup> <sup>2762</sup> <sup>2763</sup> <sup>2764</sup> <sup>2765</sup> <sup>2766</sup> <sup>2767</sup> <sup>2768</sup> <sup>2769</sup> <sup>2770</sup> <sup>2771</sup> <sup>2772</sup> <sup>2773</sup> <sup>2774</sup> <sup>2775</sup> <sup>2776</sup> <sup>2777</sup> <sup>2778</sup> <sup>2779</sup> <sup>2780</sup> <sup>2781</sup> <sup>2782</sup> <sup>2783</sup> <sup>2784</sup> <sup>2785</sup> <sup>2786</sup> <sup>2787</sup> <sup>2788</sup> <sup>2789</sup> <sup>2790</sup> <sup>2791</sup> <sup>2792</sup> <sup>2793</sup> <sup>2794</sup> <sup>2795</sup> <sup>2796</sup> <sup>2797</sup> <sup>2798</sup> <sup>2799</sup> <sup>2800</sup> <sup>2801</sup> <sup>2802</sup> <sup>2803</sup> <sup>2804</sup> <sup>2805</sup> <sup>2806</sup> <sup>2807</sup> <sup>2808</sup> <sup>2809</sup> <sup>2810</sup> <sup>2811</sup> <sup>2812</sup> <sup>2813</sup> <sup>2814</sup> <sup>2815</sup> <sup>2816</sup> <sup>2817</sup> <sup>2818</sup> <sup>2819</sup> <sup>2820</sup> <sup>2821</sup> <sup>2822</sup> <sup>2823</sup> <sup>2824</sup> <sup>2825</sup> <sup>2826</sup> <sup>2827</sup> <sup>2828</sup> <sup>2829</sup> <sup>2830</sup> <sup>2831</sup> <sup>2832</sup> <sup>2833</sup> <sup>2834</sup> <sup>2835</sup> <sup>2836</sup> <sup>2837</sup> <sup>2838</sup> <sup>2839</sup> <sup>2840</sup> <sup>2841</sup> <sup>2842</sup> <sup>2843</sup> <sup>2844</sup> <sup>2845</sup> <sup>2846</sup> <sup>2847</sup> <sup>2848</sup> <sup>2849</sup> <sup>2850</sup> <sup>2851</sup> <sup>2852</sup> <sup>2853</sup> <sup>2854</sup> <sup>2855</sup> <sup>2856</sup> <sup>2857</sup> <sup>2858</sup> <sup>2859</sup> <sup>2860</sup> <sup>2861</sup> <sup>2862</sup> <sup>2863</sup> <sup>2864</sup> <sup>2865</sup> <sup>2866</sup> <sup>2867</sup> <sup>2868</sup> <sup>2869</sup> <sup>2870</sup> <sup>2871</sup> <sup>2872</sup> <sup>2873</sup> <sup>2874</sup> <sup>2875</sup> <sup>2876</sup> <sup>2877</sup> <sup>2878</sup> <sup>2879</sup> <sup>2880</sup> <sup>2881</sup> <sup>2882</sup> <sup>2883</sup> <sup>2884</sup> <sup>2885</sup> <sup>2886</sup> <sup>2887</sup> <sup>2888</sup> <sup>2889</sup> <sup>2890</sup> <sup>2891</sup> <sup>2892</sup> <sup>2893</sup> <sup>2894</sup> <sup>2895</sup> <sup>2896</sup> <sup>2897</sup> <sup>2898</sup> <sup>2899</sup> <sup>2900</sup> <sup>2901</sup> <sup>2902</sup> <sup>2903</sup> <sup>2904</sup> <sup>2905</sup> <sup>2906</sup> <sup>2907</sup> <sup>2908</sup> <sup>2909</sup> <sup>2910</sup> <sup>2911</sup> <sup>2912</sup> <sup>2913</sup> <sup>2914</sup> <sup>2915</sup> <sup>2916</sup> <sup>2917</sup> <sup>2918</sup> <sup>2919</sup> <sup>2920</sup> <sup>2921</sup> <sup>2922</sup> <sup>2923</sup> <sup>2924</sup> <sup>2925</sup> <sup>2926</sup> <sup>2927</sup> <sup>2928</sup> <sup>2929</sup> <sup>2930</sup> <sup>2931</sup> <sup>2932</sup> <sup>2933</sup> <sup>2934</sup> <sup>2935</sup> <sup>2936</sup> <sup>2937</sup> <sup>2938</sup> <sup>2939</sup> <sup>2940</sup> <sup>2941</sup> <sup>2942</sup> <sup>2943</sup> <sup>2944</sup> <sup>2945</sup> <sup>2946</sup> <sup>2947</sup> <sup>2948</sup> <sup>2949</sup> <sup>2950</sup> <sup>2951</sup> <sup>2952</sup> <sup>2953</sup> <sup>2954</sup> <sup>2955</sup> <sup>2956</sup> <sup>2957</sup> <sup>2958</sup> <sup>2959</sup> <sup>2960</sup> <sup>2961</sup> <sup>2962</sup> <sup>2963</sup> <sup>2964</sup> <sup>2965</sup> <sup>2966</sup> <sup>2967</sup> <sup>2968</sup> <sup>2969</sup> <sup>2970</sup> <sup>2971</sup> <sup>2972</sup> <sup>2973</sup> <sup>2974</sup> <sup>2975</sup> <sup>2976</sup> <sup>2977</sup> <sup>2978</sup> <sup>2979</sup> <sup>2980</sup> <sup>2981</sup> <sup>2982</sup> <sup>2983</sup> <sup>2984</sup> <sup>2985</sup> <sup>2986</sup> <sup>2987</sup> <sup>2988</sup> <sup>2989</sup> <sup>2990</sup> <sup>2991</sup> <sup>2992</sup> <sup>2993</sup> <sup>2994</sup> <sup>2995</sup> <sup>2996</sup> <sup>2997</sup> <sup>2998</sup> <sup>2999</sup> <sup>3000</sup> <sup>3001</sup> <sup>3002</sup> <sup>3003</sup> <sup>3004</sup> <sup>3005</sup> <sup>3006</sup> <sup>3007</sup> <sup>3008</sup> <sup>3009</sup> <sup>3010</sup> <sup>3011</sup> <sup>3012</sup> <sup>3013</sup> <sup>3014</sup> <sup>3015</sup> <sup>3016</sup> <sup>3017</sup> <sup>3018</sup> <sup>3019</sup> <sup>3020</sup> <sup>3021</sup> <sup>3022</sup> <sup>3023</sup> <sup>3024</sup> <sup>3025</sup> <sup>3026</sup> <sup>3027</sup> <sup>3028</sup> <sup>3029</sup> <sup>3030</sup> <sup>3031</sup> <sup>3032</sup> <sup>3033</sup> <sup>3034</sup> <sup>3035</sup> <sup>3036</sup> <sup>3037</sup> <sup>3038</sup> <sup>3039</sup> <sup>3040</sup> <sup>3041</sup> <sup>3042</sup> <sup>3043</sup> <sup>3044</sup> <sup>3045</sup> <sup>3046</sup> <sup>3047</sup> <sup>3048</sup> <sup>3049</sup> <sup>3050</sup> <sup>3051</sup> <sup>3052</sup> <sup>3053</sup> <sup>3054</sup> <sup>3055</sup> <sup>3056</sup> <sup>3057</sup> <sup>3058</sup> <sup>3059</sup> <sup>3060</sup> <sup>3061</sup> <sup>3062</sup> <sup>3063</sup> <sup>3064</sup> <sup>3065</sup> <sup>3066</sup> <sup>3067</sup> <sup>3068</sup> <sup>3069</sup> <sup>3070</sup> <sup>3071</sup> <sup>3072</sup> <sup>3073</sup> <sup>3074</sup> <sup>3075</sup> <sup>3076</sup> <sup>3077</sup> <sup>3078</sup> <sup>3079</sup> <sup>3080</sup> <sup>3081</sup> <sup>3082</sup> <sup>3083</sup> <sup>3084</sup> <sup>3085</sup> <sup>3086</sup> <sup>3087</sup> <sup>3088</sup> <sup>3089</sup> <sup>3090</sup> <sup>3091</sup> <sup>3092</sup> <sup>3093</sup> <sup>3094</sup> <sup>3095</sup> <sup>3096</sup> <sup>3097</sup> <sup>3098</sup> <sup>3099</sup> <sup>3100</sup> <sup>3101</sup> <sup>3102</sup> <sup>3103</sup> <sup>3104</sup> <sup>3105</sup> <sup>3106</sup> <sup>3107</sup> <sup>3108</sup> <sup>3109</sup> <sup>3110</sup> <sup>3111</sup> <sup>3112</sup> <sup>3113</sup> <sup>3114</sup> <sup>3115</sup> <sup>3116</sup> <sup>3117</sup> <sup>3118</sup> <sup>3119</sup> <sup>3120</sup> <sup>3121</sup> <sup>3122</sup> <sup>3123</sup> <sup>3124</sup> <sup>3125</sup> <sup>3126</sup> <sup>3127</sup> <sup>3128</sup> <sup>3129</sup> <sup>3130</sup> <sup>3131</sup> <sup>3132</sup> <sup>3133</sup> <sup>3134</sup> <sup>3135</sup> <sup>3136</sup> <sup>3137</sup> <sup>3138</sup> <sup>3139</sup> <sup>3140</sup> <sup>3141</sup> <sup>3142</sup> <sup>3143</sup> <sup>3144</sup> <sup>3145</sup> <sup>3146</sup> <sup>3147</sup> <sup>3148</sup> <sup>3149</sup> <sup>3150</sup> <sup>3151</sup> <sup>3152</sup> <sup>3153</sup> <sup>3154</sup> <sup>3155</sup> <sup>3156</sup> <sup>3157</sup> <sup>3158</sup> <sup>3159</sup> <sup>3160</sup> <sup>3161</sup> <

copied into the own script.

Printing the annotations of several users poses the same challenges as on the screen because annotations of different users might overlap. However, in contrast to screens, a dynamic expansion of collapsed annotations is not possible. For this reason, the print module generates a layout in which the document page is scaled down. Of each set of overlapping annotations, only one annotation is printed at its correct context position, while all other annotations are moved to the margins. Their context position is marked by a connecting line (see Fig. 5.19).

**Scalability of multi-user visualizations.** An important issue of views that integrate annotations from all users is their scalability to a large number of annotations and users. An evaluation with annotations made by students in real lectures (which is presented in more details in Section 7.2) supports the assumption that in a lecture scenario, the views scale well to a larger user community. Even in very large audiences, the average number of shared annotations per slide remains rather small.

Scalability clearly does not depend on own annotations. Similarly, annotations which are made by members of the user's group do not have an influence, as the average number of members in a group is not affected by the size of the entire audience. Hence, scalability only depends on public annotations. Our experiences show that in a lecture setting, only a very small fraction of annotations is published to the whole community. In our case this were 1.6 % of all annotations. In our lecture evaluation, each student made an average of 0.59 annotations per slide. Assuming four members in a learning group, an average of 1.7 additional annotations are shared by these members. Public annotations of the entire audience average out at 0.9 annotations for 100 participants and at 4.7 annotations for 500 participants. Hence, the total number of slides keeps small.

There are however some slides which are heavily annotated. In our evaluation, the most frequently annotated slide contains an average of 5.2 annotations per user. This "worst case" results in an average of 15.2 annotations shared by the own group and an additional 8 annotations or 41 annotations for 100 and 500 participants, respectively. These extremely high numbers apply only to a very small number of slides (the top 10 % of slides have more than an average of 2.8 annotations per user). In order to cope in these situations with a too large number of public annotations, these annotations could be automatically filtered. While personal and shared group annotations are visualized as discussed above, only these public annotations that have been implicitly classified as relevant by members of the author's group are displayed to the entire audience.

### 5.3.5 Paper-based Presentations

**Scenario 13** (Annotations of the Instructor). *During the lecture, many instructors make annotations on the course script. For this purpose, they use overhead transparencies and an overhead projector or digital slides, a pen-sensitive display and a beamer. Yet, the instructor is forced to remain at one fix position in front of the overhead projector or the pen-sensitive display. With CoScribe, the instructor can make live annotations and can control the slide presentation from any point in the lecture hall. This enables a more varied and more interactive presentation style.*

An important aspect of the ecology of lecture annotations are the annotations made by the instructor during the lecture. For this type of annotations, CoScribe includes an instructor interface, which is similar to the interface first presented in [SN07]. The instructor can use a digital pen to make handwritten annotations on a paper printout of the presentation slides or on a pen-enabled display. The annotations are added to the slides and projected with the slides in real-time. Moreover, the instructor can use digital pen and paper as a remote control for the projection. By tapping with the pen on a printed slide, this slide is automatically projected.

This frees the instructor from standing at a fix position at the pen-enabled display. Instead, he or she can freely move in the lecture hall and is still able controlling the presentation and making annotations. Moreover, the instructor can lay out the individual pages of the printout and easily jump to other slides. This facilitates non-linear presentation styles which enable the instructor to be more responsive to the audience.

Our solutions improves upon [SN07] by automatically distributing the annotations of the instructor to all attendants. These can access the annotations in the document viewer in a presenter's view. Similarly to other shared annotations, relevant ones can be copied into the own script.

### 5.3.6 Handwriting Recognition

A substantial advantage of digital over traditional handwritings is the possibility to recognize the handwritten text and to offer full text search within handwritings. In this section, we will analyze the recognition accuracy for handwritten annotations and present an approach that significantly increases the performance for domain-specific terms<sup>2</sup>.

**Microsoft Handwriting Recognition Engine.** The Microsoft handwriting recognition engine is part of the Windows XP Tablet PC Edition. It is an on-line handwriting recognition engine [PS00] that recognizes text from digital ink. In contrast to off-line handwriting recognition, on-line recognition uses not only the visual image of the handwriting, but relies on spatio-temporal data, i.e. the temporal sequence of the

---

<sup>2</sup>This work was performed together with the diploma student Jie Zhou under my supervision.

two-dimensional coordinates of the writing is available and used for the recognition. The engine utilizes a built-in dictionary of words of a given language, which can be extended or replaced by other dictionaries. It automatically segments a set of pen strokes into individual words and separates text from graphics. For each word, a best guess and up to nine alternates with lower confidence scores are returned. The version available at the time of our experiments was not trainable to an individual user's handwriting. This is possible with the newest version of the engine, which is part of Microsoft Windows Vista. The engine is accessed with an SDK from the .NET platform.

**Metrics.** Common metrics used to evaluate the performance of speech recognition and handwriting recognition engines are word error rate and character error rate. Both metrics are based on the Levenshtein Distance [Lev66]. This calculates the minimal distance between two strings by examining the minimal number of operations that is required to transform one string into the other. An operation can be an insertion, a deletion or a substitution of one single character.

The character error rate is calculated as follows:

$$CER = \frac{I + S + D}{N} * 100$$

where

- I is the number of inserted characters
- S is the number of substituted characters
- D is the number of deleted characters
- N is the maximum number of characters of both text strings

The word error rate (*WER*) uses the same formula, in which however the parameters *I*, *S* and *D* refer to operations on entire *words* and *N* is the maximum number of *words*.

**Baseline performance.** We first evaluated the baseline performance of the handwriting recognition engine. From a set of 679 annotations that 10 students had made on lecture slides of a computer science lecture, we removed annotations that contain only drawings, whose text varies extremely in size (characters have more than 100 % of difference in height) or which are illegible to human readers. Of this reduced set, we randomly chose 169 handwritten annotations. The text of these annotations was manually transcribed.

As shown in Table 5.3, the resulting word and character error rates elevated at 49.2 % and 18.6 % respectively. These error rates are even more elevated than the unsatisfactory rates reported on rather similar handwritten data: Koile et al. [KCR<sup>+</sup>07]

|                      | Percentage |
|----------------------|------------|
| Word error rate      | 49.2 %     |
| Character error rate | 18.6 %     |

Table 5.3: Baseline performance of the handwriting recognition of lecture annotations

### Aufbau des Heap

- Erfüllen der Heapeigenschaft durch Vertauschen (von den Blättern aus):
  - Gehe im Array/Baum von hinten nach vorne
  - Stelle Heapeigenschaft zwischen aktuellem Knoten und direktem Nachfolger her (eventuell "einsinken lassen" -> solange weiter mit größtem Nachfolger tauschen, bis beide Nachfolger kleiner sind)

```

for (j = L.length; j >= 1; j--)
  einsinken (L [j])
einsinken:
while (L [j] kleiner als ein Nachfolger)
  tausche L [j] mit größtem Nachfolger
  deklariere die entsprechende Nachfolgerposition als i

```

Knoten werden solange "einsinken gelassen" bis Heapbedingung nicht mehr verletzt.

position als i

for (i = L.length; i >= 1; i--)  
einsinken (L [i]);

einsinken: while (L [i] < als ein Nachfolger)  
tausche L [i] mit größtem Nachfolger

Figure 5.20: Example of an annotation with text which is difficult to recognize

used the same engine and reported a word error rate of 27 % on handwritten answers that students made on Tablet PCs as a response to specific questions of the teacher.

Our goal was to find out why in our case the error rate is that elevated. For this reason, we manually analyzed all 679 annotations that were made by 10 users in this lecture. The analysis showed that these annotations are much more complex to recognize than handwritten notes or answers. In contrast to answers that are submitted to the teacher, annotations serve a personal use. They have an informal character and are often written in a hurry. The annotations we analyzed heavily varied in size, position and orientation. Even within one single annotation, the size of the characters can be very different (Fig. 5.20). Moreover, a considerable number of annotations (29 % of our test set) contained mixed text and drawings (Fig. 5.21). This is also due to the topic of the lecture – sorting algorithms – in which many tree-like structures were drawn by the students. Finally, annotations contain many (often personal) abbreviations as well as domain-specific terms (like  $O(n \log n)$ ) or formulae.

### Approaches for a Better Performance of Handwriting Recognition

Given these problematic aspects, several approaches seem helpful for increasing the recognition performance. One could examine methods for improving the module that normalizes the size of words and characters. Another approach is improving

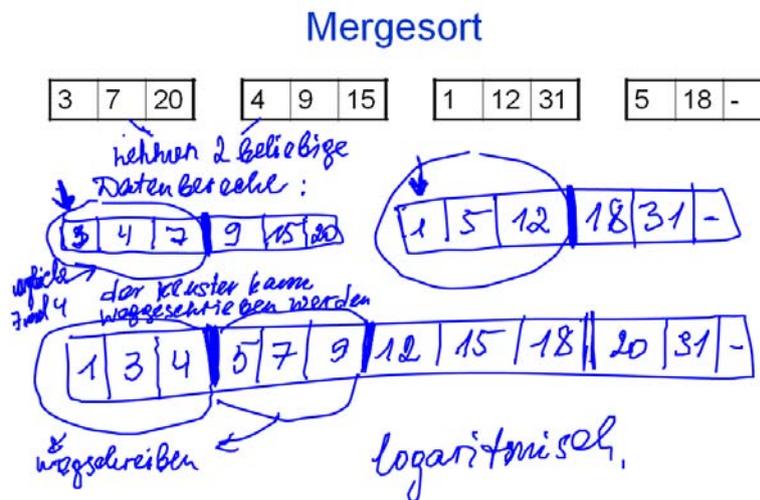


Figure 5.21: An annotation that combines text with graphics

the text/graph separation. This seems particularly viable when the application is restricted to a specific domain. For example, Koile et al. [KCR<sup>+</sup>07] added a new module for separating text from arrows, as the answers of the students in that domain contained many arrows.

A third approach – the one we examined – aims at improving the recognition accuracy of domain-specific terms. This is important because domain-specific terms are often used in search queries as well as for tagging and indexing documents.

For these experiments, we removed annotations that contain drawings from our test set. This results in a total of 118 purely textual annotations. A considerable percentage of 11.3 % of all terms in these annotations are domain-specific terms. The selection of domain-specific words<sup>3</sup> from the field of computer science and the classification of tokens was made by a domain expert (a fifth-year computer science student).

We examined three approaches, all adapting the dictionary that is used by the handwriting recognition engine. They are based on the assumption, that annotations made on lecture slides frequently contain words that are also written on these slides. In particular, this takes into account domain-specific words, which are typi-

<sup>3</sup>The following terms, contained in the slides and/or the annotations, were classified as domain-specific: Pivot, Pivotelement, Worst Case, Sortiert, Quicksort,  $O(n \log n)$ ,  $O(n)$ ,  $O(n^2)$ , Zeiger, Daten, Baumstruktur, AVL-Baum, Schlüssel, Sortierung, Zwischenordnung, Durchläufe, Durchlauf, Speicherbedarf, Speicher, Binärbaum, Array, Suchbaum, Wurzel, Knoten, Sortieren, Sortierschritte, Mergesort, Heap, Heapbedingung, Heapsort, Laufzeit, Aufwand, Einsinken, tauschen, austauschen, Liste, Teilliste, optimieren, Sortieralgorithmus, Element, Java, Bottom-Up-Verfahren, Datenblöcke, Treesort, Out-of-Place, Sortierverfahren, Permutation, Datenmengen, Cluster, In-ordertraversierung, Pseudocode, Datenstruktur, Rekursion, binär, Logarithmus, Heapbereich, rekursiv.

|                                 | Word error rate (%) | Character error rate (%) |
|---------------------------------|---------------------|--------------------------|
| No domain-specific dictionary   | 45.3 %              | 18.2 %                   |
| Dictionary A (all slides)       | 41.2 %              | 16.4 %                   |
| Dictionaries B (current slide)  | <b>36.5 %</b>       | 16.2 %                   |
| Dictionaries C (sliding window) | 41.8 %              | 17.4 %                   |

Table 5.4: Performance of the handwriting recognition for domain-specific terms

cally not included in the standard dictionary.

**Approach 1.** The first experiment adds the tokens from all slides of the given lecture to the dictionary. Hence, the dictionary is the same for all annotations from our corpus. It contains 2283 words taken from the 42 slides of the lecture. We will refer to this dictionary as Dictionary A.

**Approach 2.** The second experiment uses different dictionaries for annotations made on different slides. The dictionary for a specific annotation contains all tokens extracted from the slide the annotation is located on. We will refer to these dictionaries as Dictionaries B.

**Approach 3.** The third experiment relies on a sliding-window approach. Again, annotations made on different slides have different dictionaries. The dictionary for a specific annotation contains all tokens extracted from the slide the annotation is located on and all tokens from the preceding five and the following five slides. If the slide is amongst the first or last five slides, the smaller number of all preceding or all subsequent slides is used. We experimented with different numbers of preceding and subsequent slides and found that in our case, the number of five slides provides the best results. As a matter of course, this number depends on the specific slide set. We refer to these dictionaries as Dictionaries C.

**Results.** Table 5.4 gives the recognition results for domain-specific words for all three types of dictionaries and contrasts them with the result if no domain-specific dictionary is used. The use of either dictionary significantly reduces the word error rate. Dictionary B (all tokens from the current slide) clearly outperforms the other dictionaries. In contrast to using no domain-specific dictionary, the word error rate could be reduced by almost 20 %.

**Implications for the design.** The results of this evaluation show that handwritten annotations cannot be recognized with a high accuracy using a standard handwriting recognition engine. Due to this low recognition accuracy, the interaction design of CoScribe does not require handwriting recognition. In particular, handwriting is

not displayed as text but as the original handwriting. Nevertheless, handwriting recognition is used in the background. The text of each newly created or modified annotation is automatically recognized. We have seen that the recognition performance for domain-specific terms can be significantly improved by using domain-specific dictionaries, which are automatically created from the lecture slides. The recognized data is then used for full-text search within the annotated slides. Moreover, handwriting recognition is optionally used for recognizing handwritten labels of tags. These typically do not involve the problems of annotations, which were described above, because labels are typically written more accurately, do not include drawings and have a fix size and orientation of text.

## 5.4 Linking Documents

In the previous section, we have presented interaction techniques and visualizations for shared handwritten annotations. In this section, we will discuss a more formal type of annotations, that being references created by the user. The following scenarios give some examples of situations in which references support learning.

**Scenario 14** (Integrating Information from Many Documents). *In addition to her lectures, Sally attends a seminar. Her task is to compose a term paper on a specific topic. This requires mainly autonomous, self-directed work. The instructor indicates a few pointers to the literature. Based on these, she has to find further literature and precise the subject and the formulation of the questions. She searches literature in libraries, borrows books and photocopies some journal articles. She also looks for information on the web and prints web pages and PDF documents in order to have them available on paper. Initially this is a rather chaotic and unstructured collection of various documents. In order to structure the collection, Sally connects related documents by hyperlinks. A graphical overview on all documents and the hyperlinks between them is automatically created by CoScribe.*

**Scenario 15** (Excerpting). *Excerpting is an important scientific method for reading books and summarizing their contents in a structured way. An excerpt is a short summary of the most important aspects of an existing text. This can also include own thoughts of the reader, can set the text into a larger context and can establish relations with previously acquired knowledge. During reading a text, the reader makes notes on a separate document, e.g. a paper notebook containing empty pages. He or she adds references to the original text to indicate the passage a note refers to. This allows to find the original passage later on.*

Our field studies have shown that learners frequently work with an interwoven set of documents, where information is distributed across several documents (e.g. in a printed lecture script, in some additional Web pages and in a personal notebook). In this case, references are an important means that help the user in integrating this information. Traditional handwritten references are difficult to follow, since the target document must be manually searched. This is particularly hard with cross-media hyperlinks. Moreover, users cannot easily share their references with co-workers. And even if references of co-workers are shared, it is not easy to understand their meaning because users have highly individual practices of referencing.

CoScribe supports users in adding own hyperlinks to existing documents in order to connect printed and digital documents in any combination. These cross-media hyperlinks are not only practical, being faster and easier to follow, but the conscious selection of passages to be linked supports successful learning. By abstracting and establishing relationships, learners build structural knowledge, which facilitates recall and comprehension and is essential to problem solving [JBY93].

This chapter is organized as follows. After defining the model of links, we will present the interaction concepts for creating and following hyperlinks. Finally, we

| Desideratum  | Approach  | Section | Concept | Function | Innovation w.r.t. related work | Main re-quirements    |
|--|---|---------|---------|----------|--------------------------------|-----------------------|
| Easy and reliable creation of cross-media hyperlinks | Association gesture spanning physical and digital documents | 5.4.2   | •       | •        | •                              | R1, R1a, R2, R4a, R4c |
| Flexible scopes                                      | Different types of association areas                        | 5.4.3   | •       |          | •                              | R4, R4a               |
| Integration of legacy media                          | Physical folders  |         | •       |          | •                              |                       |
|  | Anoto-enabled registration stickers                         | 5.4.3   | •       | •        | •                              | R4                    |
| Quick and easy following of hyperlinks               | Tapping on link hot-spot                                    | 5.4.4   | •       | •        |                                | R4a                   |
|  | Same interaction in both worlds                             |         | •       |          | •                              | R1a                   |
|  | Cross-media history   |         | •       | •        | •                              | R1                    |
| Sharing hyperlinks                                   | Button toolbar  | 5.4.5   | •       | •        | •                              | R1, R3, R4a           |
| Access to shared hyperlinks                          | Integrated multi-user visualizations                        | 5.3.4   | •       | •        | •                              | R1, R3, R4a           |
|  | Ecological view   | 5.2.5   | •       | •        | •                              | R1, R3, R4c           |

Table 5.5: Desiderata and our approaches concerning cross-media hyperlinks

will discuss how hyperlinks can be shared and how shared links are visualized. Table 5.5 provides an overview on the challenges which we address in this chapter.

### 5.4.1 Model of Links

A hyperlink is modeled as a binary association between two document entities. It is symmetric and can be traversed in both directions. These two-way links follow the original hypertext principles introduced by Bush [Bus45] and Nelson [Nel65] and stand in contrast to the reduced model of one-way hyperlinks known from the World Wide Web.

It is established practice to create references between passages of different extent. This includes referencing an individual figure or a short paragraph, entire pages or

chapters, a whole book or even several documents. To account for this practice, our model supports the entities to be linked to be of different extent. An entity may have either of the following scopes:

- a collection of documents,
- an entire document, or
- an area within a document. This area can span several columns and/or pages.

Each document or document collection can be represented on paper or on the screen. Hence, associations can span the paper/digital boundary.

While we model passages within documents as regions of space, our interaction technique could be coupled with the automatic extraction of document elements [WISN08]. This would allow linking specific semantic objects within the document.

### 5.4.2 Association Gestures

**Underlying assumptions.** In the early design phase, we investigated how students and colleagues make traditional handwritten references in documents. These are typically not made within the text but in the left or right margin of the document, leading to clearly visible references. Inspired by this practice, we opted for a spatial differentiation between the actual document and association areas. In these rectangular areas besides and above the document, users can create and follow hyperlinks. Association areas are both printed on paper documents and displayed in digital documents.

A key issue of our interaction design is the separation between a generic association gesture and different association areas, which act as end points for the gesture and represent different scopes. This combines intuitive interaction with versatility and predictability even with the restricted feedback capabilities of a Pen-and-Paper User Interface.

In the following, we first present the pen gesture for creating hyperlinks and then provide more details on the different types of association areas this gesture is performed on.

Fundamentally different techniques for creating hyperlinks are possible. One could create hyperlinks by physically moving different documents or sheets of paper that shall be linked. For example, these could be shortly placed on top of each other to create the association. A second approach could consist of performing specific gestures with the fingers or with the hand, for example simultaneously placing both hands on the two documents that shall be linked. Instead of using hands or fingers, gestures could also be made with a digital pen. A third approach consists of manipulating the physical shape of the documents, e.g. by temporarily bending or folding them or by attaching two stickers on both anchors of the link as proposed by [HHT99].

Our approach relies on pen gestures for several reasons. First, the user makes annotations with a digital pen. Using the same pen for linking tightly integrates both

interactions. Second, in contrast to gestures performed with fingers or hands or manipulations of the physical shape of a document, no further tracking infrastructure than the digital pen – which is already used for annotations – is necessary. While at a first glance, multi-touch seems compelling for associating two link anchors in one single step, our experiences showed that users typically first select a first link anchor and then seek for the appropriate second position of the association. This is better served by an interaction which is divided into two steps (and hence is possible with one single pen). Finally, pen gestures are very flexible. It is possible to mark and link small passages, but also entire documents. Pen gestures can be made both on paper and on pen-enabled displays. In contrast, linking documents by physically moving them is intuitive for entire documents but not for flexible sub-passages within a document. Moreover, moving or collocating digital documents requires a different interaction than moving or collocating physical documents.

**The pen gestures.** Similar to the interaction technique for annotating documents, our technique for creating hyperlinks is designed to be minimally invasive and highly compatible with existing practices of creating references. The starting point are informal handwritten references (e.g. “see p. 4”, “cf. Wikipedia Digital Paper”), which indicate the link target in a way that can be interpreted by a human reader. This informal way of referencing is maintained. In order to additionally enable interpretation by computers, the user performs a quick, simple and reliable pen gesture. Accounting for the associative nature of hyperlinks, this is an association gesture which connects two association areas.

Since one of our key design aspects is supporting cross-media linking, this gesture typically spans two different physical documents or associates a physical document and a digital document on a display. Our interaction model considers both a paper page and a screen as similar display instances. This interaction is therefore an extension to prior stitching gestures that span different paper documents [Rek97] or several displays [HRG<sup>+</sup>04].

Using one single pen, this gesture is performed on paper, on the pen-sensitive display or on both of them. There exist two variants of the pen gesture. Figure 5.22 gives a visual overview on the gestures. Figure 5.23 provides details on the gestures.

- *Single-line gesture:* In order to define the first link anchor, the pen is held down on a first association area for 500 ms without moving (until a click sound is played). Without lifting the pen, the user then draws a line to the second association area, where she finally lifts the pen (Fig. 5.22 (a) and Fig. 5.23 lower left).
- *Two-point gesture:* Alternatively, if the two areas are not close to each other, the user makes two consecutive pen taps on both link anchors instead of the connecting line (Fig. 5.22 b). This has the disadvantage that the interaction is not modeless but on the other hand, the interaction does not require repositioning documents. This is for example necessary for links within the same

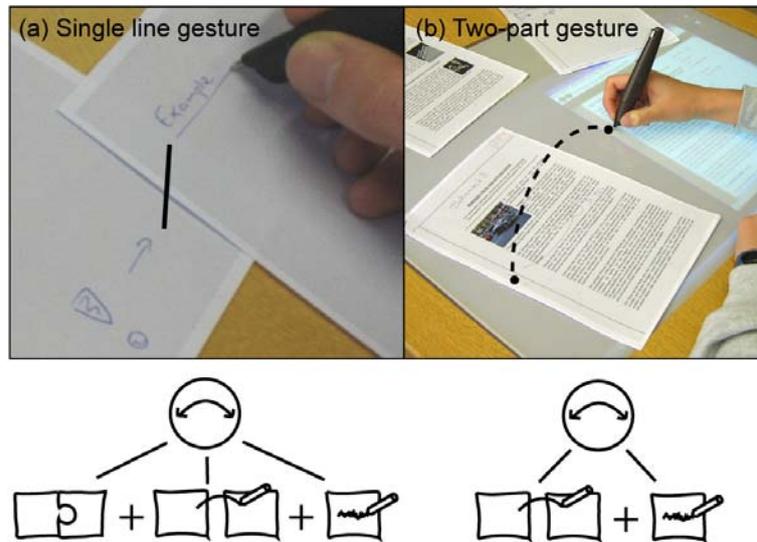


Figure 5.22: Pen gestures for creating a hyperlink

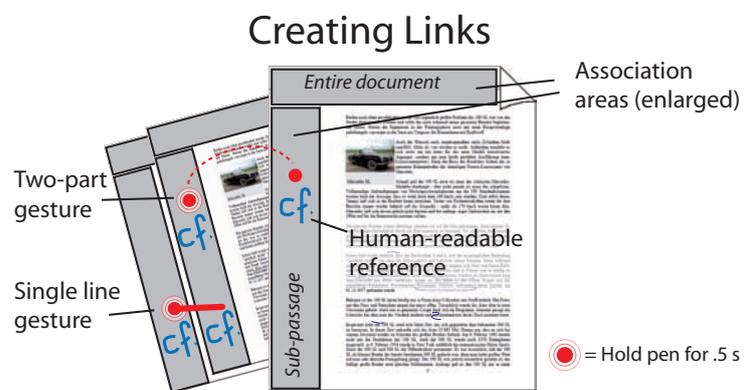


Figure 5.23: Details of the pen gestures

book. Moreover, the user can start a link already before having selected the exact link target.

The requirement of holding down the pen without moving for 500 ms permits a clear distinction between annotations and link commands. Moreover, it avoids confusing link start and link ending in the modal variant. If a computer is nearby, the gestures are recognized in real-time and instant audio feedback is given. A link can be deleted with a cross-out gesture on any marking made for creating this link.

This associative gesture is very easy and can be quickly performed during the work process.<sup>4</sup> Moreover, it does not rely on a possibly error-prone recognition of special keywords or symbols. This is particularly relevant with current pens not being able to provide system feedback.

### 5.4.3 Association Areas for Defining Scopes

This generic gesture has different meanings depending on the areas it is performed upon. This provides for an intuitive and flexible way of defining different link anchors. Different kinds of association areas are contained on printed and digital documents, on physical books and on physical folders. Each type of area represents another scope. Figure 5.24 provides an overview on these types of association areas. It shows how the physical areas are complemented by digital representations. In the document viewer, the areas are displayed at the same positions on the documents as on paper. In the ecological view, the nodes of the graph that represent the documents act as digital association areas. On these areas, the pen gesture for creating hyperlinks can be performed in the same manner as on paper. We now discuss each type of association area in turn.

- *Entire hybrid document*: Each hybrid document contains an area on the top of its first page where gestures are made for links that apply to this entire document (Fig. 5.23 left). Since one of our key design aspects is supporting cross-media linking, these association areas are included in each printed document page, in their digital representation in the document viewer and on Web pages in

---

<sup>4</sup>Related research [LHGL05] compared several pen-based techniques for switching between modes, one of them being tap-and-hold. The authors found tap-and-hold to be slower and more error-prone than switching mode by varying the pressure or by using additional buttons. We nevertheless opted for tap-and-hold for the following reasons. First, pressure-based mode switching is not possible in paper-only environments, as it is important to provide real-time feedback to the user about the current pressure level. Second, other techniques which require additional hardware were not acceptable in our case, as it heavily restricts deployability in real settings. Third, while tap-and-hold requires the user to wait a short moment before making the actual gesture, time is not critical in our case, since not so much links are created. Finally, the main problem of tap-and-hold reported was that users had difficulties in holding the pen at a fix position on the slippery display of a pen-sensitive screen. This was no problem with CoScribe, as paper is not slippery and the screen is large enough to allow some minimal movements. Our design decision was confirmed in the evaluation. Users had no difficulties in creating hyperlinks and reported to appreciate the clear distinction between a normal mode and the gesture for creating a hyperlink.

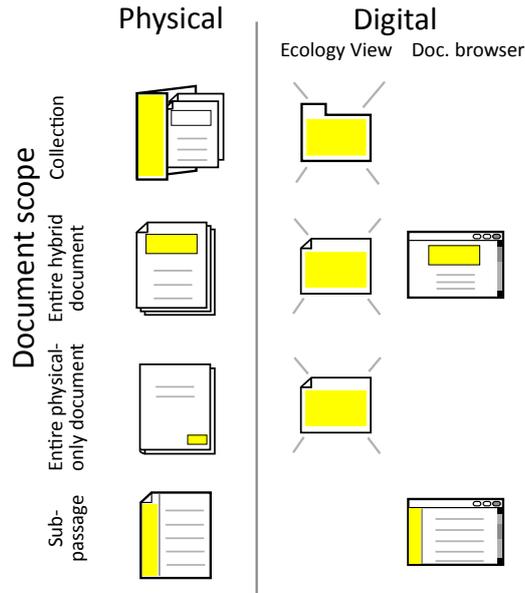


Figure 5.24: Types of association areas

Mozilla Firefox. CoScribe therefore includes a Firefox extension which manages and displays areas and pen-based hyperlinks (see Section 6.1.2 on implementation details). Another digital association area is the thumbnail of the document in the ecology visualization (see Section 5.2.5).

- *Entire physical-only document*: In order to support interactions with Anoto pens on physical-only documents (books, journals etc.), which do not contain the Anoto pattern, users can attach a small sticker covered with the pattern onto the document cover (Fig. 5.25) and register this sticker with the document's barcode. Depending on the pen technology, the user can directly scan the barcode with the digital pen or she manually enters the barcode. The digital metadata of this document (including an image of the cover, if available) is then automatically retrieved from a database. (Our implementation uses the Amazon.com web service.)
- *Sub-document level*: A margin area contained on each page of hybrid documents provides for creating links from or to sub-passages within a document. It is included both in printed documents and in the CoScribe viewer for digital documents. The user can define the precise extent of the linked document passage. She therefore draws a vertical line besides the passage where the association gesture is made. Such scope markings are deleted the same way as association gestures by a zig-zag gesture over the marking. If the document contains several columns, each column has an own association area. A scope may span several columns and/or pages.



Figure 5.25: Links from and to physical books or journals are enabled by a small sticker which is covered with Anoto pattern

In a user study on how people make references within traditional paper documents, we found that users often do not explicitly define the scope of a reference. Instead, the scope remains fuzzy and refers to an unspecified passage or item near the reference. Prior systems [LCC07, LGHH08] do not account for that. Our interaction also supports fuzzy scope definitions. These associations are made in the association column without any further specification.

In addition to these formalized gestures that are computer-interpreted, users can make any informal markings for the scope definition, which are displayed like normal handwriting in the electronic viewer (e.g. brackets, lines and arrows). While users understand them, the system is agnostic of their meaning.

- *Document collections*: For links from or to a collection of documents (one-to-many or many-to-many), CoScribe includes the concept of folders. Physically placing a document into a folder is a very intuitive traditional interaction for defining a collection of paper documents. Users can then easily reference the entire collection by referring to the folder.

Our slightly formalized interaction design is inspired by this practice. A folder (Fig. 5.26) contains one or several documents and may be the starting or ending point of an association gesture, as it contains an association area on its cover. Similarly, the digital representation of a folder in the ecological view acts as an association area.

In order to detect that physical documents are added to or removed from a folder, the interaction technique relies on location tracking of document and folder positions. At locations where no tracking infrastructure is available, the user can alternatively perform a pen gesture to inform the system about an added or removed physical document. The same gesture applies if a user wants to add or remove a digital document.

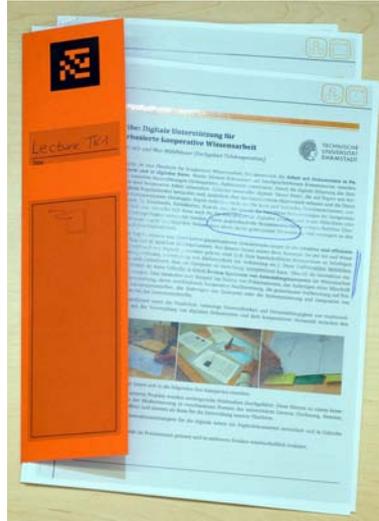


Figure 5.26: Physical folders provide for defining collections of documents in an intuitive manner

Summing up, the conceptual activity of creating links combines up to three core interactions of our framework, namely *inking*, *combining* and *bridging*. It moreover illustrates the principle of *spatial multiplexing*. The generic association gesture has different meanings depending on the area it is performed upon. Finally, *gestural multiplexing* enables to differentiate between creating associations and defining sub-document scopes.

#### 5.4.4 Following Hyperlinks

**Scenario 16** (Pre-Structured Document Collections). *Sally attends a seminar on interpretations of Molières comedies. The instructor has prepared an extensive collection of documents. This contains important passages of some comedies. Each of them is included in several editions. Moreover, secondary texts were added. The instructor connected related passages with hyperlinks. This pre-structured document collection is the basis for the interpretative work during the semester and made available in CoScribe. Sally prints the most important texts and accesses the other texts in their digital versions. Taken the printed texts as a starting point, she follows the associative trails for her interpretative activity.*

A digital hyperlink is followed by tapping on or near a link marking (the gesture or the human-readable reference) in an association area (Fig. 5.27). This is possible on a printed document page, on a digital document on the display, on a book and on a folder cover.

The target document is then displayed in the CoScribe viewer or in Mozilla Firefox. The passage the hyperlink is referring to is highlighted in yellow. If the target is a folder or a physical-only document, it is displayed in the ecological view.

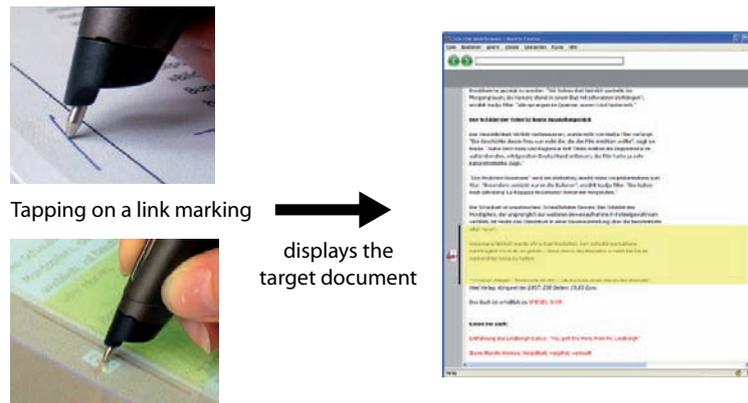


Figure 5.27: A hyperlink is activated by tapping on the link marking

Backward and forward buttons support navigating in a cross-media history that contains not only Web pages as traditionally known from Web browsers, but also the digital representations of printed documents.

It is important to note that following digital hyperlinks always leads to a digital representation, regardless if the source document is a printed or a digital one. This provides for visually skimming the contents of the target scope. If the user prefers to engage with the document more deeply (e.g. by active reading), he or she either can print a paper version or is indicated the location of an already existing printed representation (see Section 5.2.6). For this reason, it is no significant drawback that links do not directly point to physical document instances.

### 5.4.5 Collaboration

As with annotations, CoScribe supports sharing hyperlinks with other users. The default setting is that hyperlinks are visible to the other members of the user's group. If the user wants to set a different visibility, she can use classification buttons introduced for annotations in Section 5.3.3. Shared hyperlinks are digitally accessible both in document and ecological views.

In the document view, shared links of other users are displayed in a similar manner as shared annotations (see Section 5.3.4 above). In addition to the annotations of a particular user, the single-user view includes the hyperlinks made by this user. In the multi-user view, the symbols of shared hyperlinks are displayed in a scaled-down manner. Similarly to shared annotations, shared hyperlinks can be copied into the own script or can be removed from the view. The visualization of shared hyperlinks in the printed representation follows the same principle as the visualization of shared hyperlinks are visualized like shared annotations.

Shared hyperlinks are not only available in the views of individual documents, but also in the ecological view. In this view, a hyperlink is visualized by an edge

between two nodes that each represent a document, a document page or a folder (see Section 5.2.5).

## 5.5 Tagging Documents

A third important aspect besides annotating and linking documents is to support knowledge workers in categorizing and structuring the contents of the information ecology. As we have seen in Chapter 2, multiple printed and digital documents form one document space covering the problem domain. Structuring this problem domain, abstracting to higher-level concepts and establishing relationships between concepts is crucial for successful knowledge work. For this purpose, tagging documents is a highly relevant activity. Tags can transform an unsorted and possibly confusing collection of a large number of disparate documents to a unified and well-structured document space. Moreover, it is much more easy to search and find specific information in a well-structured document space.

**Scenario 17 (Abstracting).** *Sally uses tags to prioritize documents according to their relevance for her research questions. In a next step, she systematizes the information. She identifies abstract concepts and tags each important passage with the appropriate concept. By doing so, she has not only understood the individual aspects but has acquired structural knowledge of the problem domain.*

These are only two examples of how tags can be used for successful learning. Tags can for example be used for:

- Structuring a domain with abstract concepts and relating these to document passages
- Marking up the structure of a document (e.g. by tagging the section headings)
- Prioritizing documents and structuring the task (e.g. by tagging documents with the labels “Prio 1” to “Prio 5”, “important”, “unclear”, “to do”)
- Recommending documents to collaborators (e.g. by tagging documents with his or her name)
- Labeling a document with the physical place it is filed for easily finding it later on
- Structuring the process of meetings (e.g. tagging a temporal phase as “elaboration phase”)

This chapter presents pen-based and tangible interaction techniques and accompanying visualizations for cross-media tagging in a mixed physical and digital environment.

We define a tag as follows:

**Definition 13 (Tag).** *A tag is a concept which is assigned to an entity, i.e. to an artifact or to a process. The concept is represented either by a predefined category (category tagging) or by one or several keywords which are freely chosen by the user (free tagging). The same concept can be assigned to several entities. If appropriate, we will distinguish between the tag concept (a specific category or a given keyword) and the instances of this tag (the assignments of this concept to entities).*

Offering different categories has two main advantages: First, the availability of specific categories can support meta-cognitive learning processes by encouraging and reminding students to perform specific important learning activities. Second, some common semantic classes facilitate the computer interpretation of tags, the sharing with other learners and automatic aggregations. In contrast, free tagging is more flexible and can be used for a wide variety of purposes, for instance for recommending contents to a collaborator by tagging them with his or her name or for prioritizing contents.

In the remainder of this section, we will present four interaction techniques for tagging. These four techniques complement each other, serving different purposes and having different properties. The choice of an appropriate technique depends both on the task and the user's personal preferences. The techniques provide for tagging individual annotations, entire pages of documents, flexible document scopes and processes. After having presented all techniques, we will compare and contrast them. Table 5.6 provides an overview on the challenges which we address in this chapter.

### 5.5.1 Tangible Tagging with Stickers: Digital Paper Bookmarks

Paper bookmarks have proven to be an efficient means both for a quick access to specific pages or parts of a paper document and for individually structuring a document. Impressive examples are books of law students, which often contain several dozens of bookmarks stuck on the margins of the pages (Fig. 5.28 left). As a metaphor, bookmarks also quickly became familiar in the electronic world for marking documents on the Web. Empirical work shows that their use positively influences the perceived ease of finding information [TBK04]. We aim to leverage the ease of this interaction also for structuring *digital* documents via their printouts.

With CoScribe, we introduce Digital Paper Bookmarks, which span both the paper and the electronic worlds. Digital Paper Bookmarks are Post-it like adhesive stickers of different colors which can be attached to physical pages of printed documents at arbitrary positions (Fig. 5.28 middle and right). They are covered with Anoto pattern and can therefore be labeled with a title using an electronic pen. They are synchronized with the digital system and serve as digital bookmarks for these pages. Hence, Digital Paper Bookmarks combine the advantages of intuitive paper-based bookmarking with digital support. For example, bookmarks can be used for tagging the structure of a document by bookmarking the beginning of each section or for marking important passages of a document.

| Desideratum  | Approach  | Section              | Concept             | Function | Innovation w.r.t. related work | Main requirements |               |
|--|---|----------------------|---------------------|----------|--------------------------------|-------------------|---------------|
| Free tagging   | Digital Bookmarks, Menu Card                                  | Paper Tag            | 5.5.1, 5.5.2        | •        | •                              | •                 | R2, R4c       |
| Category tagging   | Digital Bookmarks, Button Tagging, Process Cube, Process Knob | Paper Tagging, Cube, | 5.5.1, 5.5.3, 5.5.4 | •        | •                              | •                 | R2, R4c       |
| Full visual feedback on paper; easy physical access                          | Tagging with paper stickers                                   |                      | 5.5.1               | •        |                                | •                 | R1, R4a       |
| Collaborative structuring of documents & meta-cognitive support              | Digital Bookmarks   | Paper                | 5.5.1               | •        | •                              | •                 | R1, R4a, R4c  |
| Tagging with flexible scopes; convergence of tags; operations on the tag set | Defining tags on a separate paper tool                        |                      | 5.5.2               | •        |                                | •                 | R1a, R4a, R4b |
|  | Tag Menu Card   |                      |                     |          | •                              | •                 | R1a, R4a, R4b |
| Quick tagging while annotating   | Button Tagging  |                      | 5.5.3               | •        | •                              | •                 | R1, R4a       |
| Tagging of temporal processes  | Manipulation of tangible objects                              |                      | 5.5.4               | •        |                                | •                 | R3, R4c       |
|  | Exemplary model of phases                                     |                      |                     | •        |                                | •                 | R4c           |
|  | Process Cube  |                      |                     |          | •                              | •                 | R3, R4c       |
|  | Process Knob  |                      |                     |          | •                              | •                 | R3, R4c       |

Table 5.6: Desiderata and our approaches concerning cross-media hyperlinks



Figure 5.28: Paper stickers are a powerful means for indexing documents (left). Digital Paper Bookmarks (center and right) take on this principle and serve as tangible and digital bookmarks

### Interaction Technique

Users can easily create Digital Paper Bookmarks via three steps (see Fig.5.29):

1. *Combining*: Bookmarks can be attached to arbitrary pages of the printed document at arbitrary positions. Several bookmarks may be attached to the same page.
2. *Associate (bridging)*: With a simple pen gesture, the paper bookmark can be associated with the page it is attached to. It is then also available as digital bookmark. This association gesture is a short line connecting the bookmark with the page it is attached to. An arrow is printed on the bookmark to indicate where to draw the line (see Fig. 5.30).
3. *Label (inking)*: Since the Anoto dot pattern is printed on the stickers, one can use the electronic pen to write a tag or title on them, which will be synchronized.

The interaction thus combines three core interactions of our interaction model (see Section 4.2). The association step is not necessary if technology for tracking the location of paper sheets is used, which automatically detects the combination. The manual association lowers the technical requirements and enables bookmarking during mobile use. This proposed solution for creating bookmarks is entirely mobile and highly compatible with existing practices.

Once created, Digital Paper Bookmarks can be easily altered:

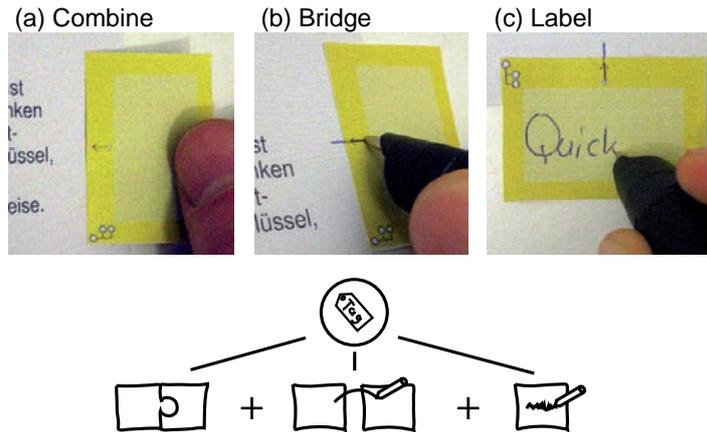


Figure 5.29: Three intuitive steps allow to create a Digital Paper Bookmark. These correspond to three core interactions of our model

- *Modify*: The position of a bookmark can be modified by sticking it to another position and performing the association gesture at this new position. Handwritten labels can be modified with the digital pen later on as well.
- *Delete*: Bookmarks can be deleted by physically removing them from the paper sheet and writing a cross out deleting gesture on the bookmark.

Since the bookmark is visible on paper, full feedback is available without additional digital support. Moreover, its shape provides a strong affordance for quickly accessing the physical page which has been bookmarked. This is a clear advantage over other classification means discussed in the literature (e.g. [LGHH08, WSS97]), which do not modify the physical shape of the document.

However, it is precisely this advantage that makes it difficult to re-print a document including bookmarks. In order to address this problem with current technology, the printout module prints a small image of the bookmark at each position of the document where a bookmark was attached in the original document. This is a visual indicator for the user who can then add bookmarking stickers at these positions. Clearly, this is a workaround which is not very satisfactory. Future printers could solve this problem by automatically attaching bookmarking stickers at the appropriate positions while a document is printed.

### Semantic Types

In addition to free tagging with keywords, Digital Paper Bookmarks offer predefined categories that support users in structuring documents. Users can choose between different types, each represented by a specific color and a specific symbol.

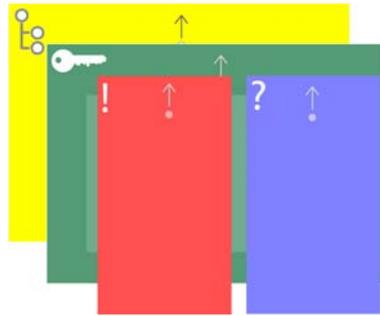


Figure 5.30: The current version of CoScribe includes four semantic types of bookmarks

Offering different types of bookmarks has two main advantages: First, some common semantic classes facilitate the computer interpretation of bookmarks, the sharing with other users and automatic aggregations. Second, if bookmarks are used in a learning setting, the availability of specific types can support meta-cognitive learning processes by encouraging and reminding users to perform specific important learning activities which are related to bookmarking.

We developed a first set of semantic types to be used with Digital Paper Bookmarks in learning processes. It comprises four types. We distinguish the following four types on two layers:

The two semantic types on the *structure layer* support learners in structuring the learning documents:

- Section bookmarks (yellow) mark the beginning of a new section. The name of this section can be written on the bookmark.
- Key bookmarks (green) can be used to mark key passages, where subject matters are defined or explained.

The two semantic types on the meta-cognitive layer serve for planning further learning activities as well as controlling past activities. They can be used for assessing both the relevance of subject matters and own difficulties.

- Important bookmarks (red) mark passages which seem particularly relevant to the learner.
- Unclear bookmarks (blue) point out passages which the learners have difficulties understanding and which therefore require further work. Once the subject matter is well understood, the bookmark can be removed.

The bookmarks of the structure layer offer much space for writing a title, whereas the meta-cognitive bookmarks serve as simple category markers without a title (Fig. 5.30).

We assume that the bookmarking actions keep intuitive and simple albeit the certain amount of abstraction imposed on the users by these four categories: On the one hand, marking documents with bookmarks of different colors is widespread in the paper world; on the other hand, the relatively small number of four categories still seems manageable. In order to help novel users to memorize the meanings of the different colors, small symbols are printed on the paper bookmarks.

### **Support of cross-media navigation**

Cross-media bookmarks provide several benefits over classical bookmarks. First, they combine the advantages of electronic bookmarking with the fact that paper is tangible and offers many possibilities to be arranged and combined in the physical space. This is especially helpful for navigation support since paper bookmarks jut out of the paper document. This affords quickly accessing the corresponding page with a single movement.

A second benefit of cross-media bookmarks is the support of orientation and navigation in case of cross-media work. As CoScribe relies both on the printed and the digital document instances, which each offer unique benefits depending on the specific task, users frequently switch between printed and digital representations of documents. For example, the user is likely to switch from the printed to the digital representation for viewing new shared annotations that other users made on this document. The user might switch in the reverse direction when she desires to deeply engage with the document and therefore prefers the printed representation. These switches require quickly finding the corresponding position in the other representation.

Support for the direction of finding the digital document instance which corresponds to a physical instance has been developed in other research (e.g. [LRH00, SN07]) and is also included in our system.

The reverse direction is more challenging. Even if the location of printed documents is electronically tracked and can be indicated by the system, this provides no support for finding a specific page within a document. Digital Paper Bookmarks support this specific task by offering synchronized visual anchor points within both printed and digital document representations. At a single glance, the user gets information about the approximate position of the page within the document, about the nearest bookmark and the page distance to this bookmark. She can then find the paper page by recurring to its approximate position in the document or by choosing an appropriate bookmark located near this page.

### **Visualization**

As Digital Paper Bookmarks are automatically synchronized with the electronic system, paper bookmarks can be integrated to and displayed within the corresponding digital document. CoScribe includes three different visualizations of bookmarks.



Figure 5.31: Visualization of bookmarks in the CoScribe viewer

**Personalized index of contents.** In the CoScribe viewer, bookmarks are visualized along with the document pages as they appear on paper (Fig. 5.31). Clicking or tapping on a bookmark displays the bookmarked page. Hence, using bookmarks automatically creates a personalized index of contents of the document including passage headings (yellow) as well as pointers to key passages (green), particularly relevant passages (red) and unclear passages (blue). In addition, bookmarks are also visualized in the slide preview panel.

**Visual overview for cross-media navigation.** A specific bookmark view in the software viewer displays a three-dimensional representation of the paper stack. It visually indicates the number of pages of the document, the current position within it and the bookmarks associated to it (Fig. 5.31 lower left). We designed this view with the goal to provide visual information helping to find the corresponding paper instance of the page currently displayed in the viewer. Furthermore, the bookmark panel also supports navigation within the digital document. Pages can be displayed by clicking or tapping on the three-dimensional representation of the paper sheet or on a bookmark.

We decided not to integrate shared bookmarks of other users in these visualizations, as they would lose their key property of mirroring the bookmarks of the printed representation. This could be very confusing when using bookmarks for cross-media navigation.



Figure 5.32: A collaborative view for comparing the structuring of documents. It includes an abstracted representation of the document (a), own bookmarks (b), bookmarks of other members of the user's learning group (c) and an aggregated view of the bookmarks of all users (d).

**Collaborative structuring.** Through the contrasting of their own structure with the structure of others, with their markings of relevant or unclear passages, learners can assess and improve their own understanding of the material. Cognitive conflicts may arise and can lead to a modification of one's own bookmarks. This is supported by a third, collaborative visualization which focuses on the structure of an individual document. It is depicted in Fig. 5.32.

The pages of a document are aligned vertically and represented in an abstracted manner to the left (Fig. 5.32 a). The user's own bookmarks are visualized beneath (Fig. 5.32 b). The bookmarks of one or several members of the user's learning group can be displayed at a time and are visible to the right of the own bookmarks (Fig. 5.32 c).

In addition to this within-group sharing, the bookmarking data of all users is automatically aggregated by the system. Anonymity is preserved as only the positions and types of the bookmarks are taken into consideration, but not their handwritten labels. An aggregated view visualizes data from all users indicating the type, position and frequency of bookmarks by colored markings of different sizes (Fig. 5.32 d). Large red markers indicate that these passages are judged particularly relevant by a large number of learners, green rectangles mark key passages, etc.

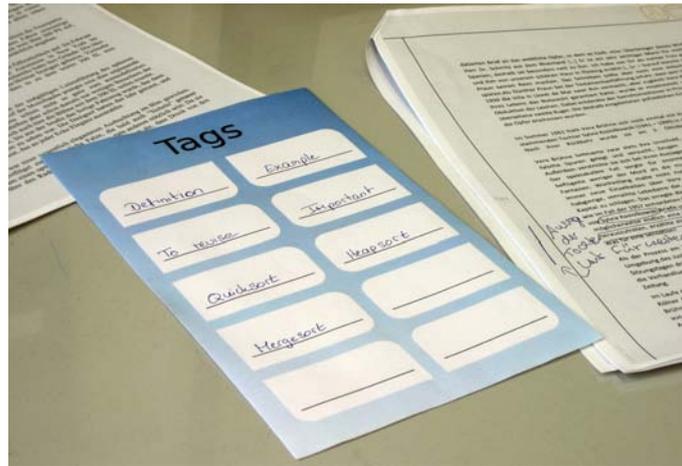


Figure 5.33: A Tag Menu Card provides for defining and applying keyword-based tags.

### 5.5.2 Tagging by Association: Tag Menu Card

While Digital Paper Bookmarks are a very intuitive interaction technique, tagging with physical stickers is rather course-grained and takes some time. Moreover, only a restricted number of bookmarks can be attached to a document without the bookmarks becoming confusing. Therefore, we introduce a second interaction strategy for tagging documents. This relies on one or more separate paper cards for defining and applying keyword-based tags (Fig. 5.33).

#### Interaction Technique

Each Tag Menu Card contains several empty areas. At any time the user may define a new tag by writing one or several freely-chosen keywords in one of these areas. After a tag is defined, it is applied using either of the following interactions:

- Writing the tag on an association area of a document (as defined in Section 5.4.2) and enclosing it with a circle in order to mark it as a tag (Fig.5.34 a). The tag is automatically recognized from the set of previously defined tags on the Tag Menu Card using handwriting recognition.
- Writing the tag on an association area and additionally performing the pen gesture for hyperlinks (Section 5.4.2) to associate it with the corresponding area on the Tag Menu Card (Fig.5.34 b). This small additional effort ensures that tagging is correctly performed, as it does not rely on handwriting recognition. This is important when no computer is nearby, as current pens cannot provide feedback on the recognition.

Although the Tag Menu Card has a printed representation, it can be used for tagging both printed and digital documents. The precise scope of the tag within the

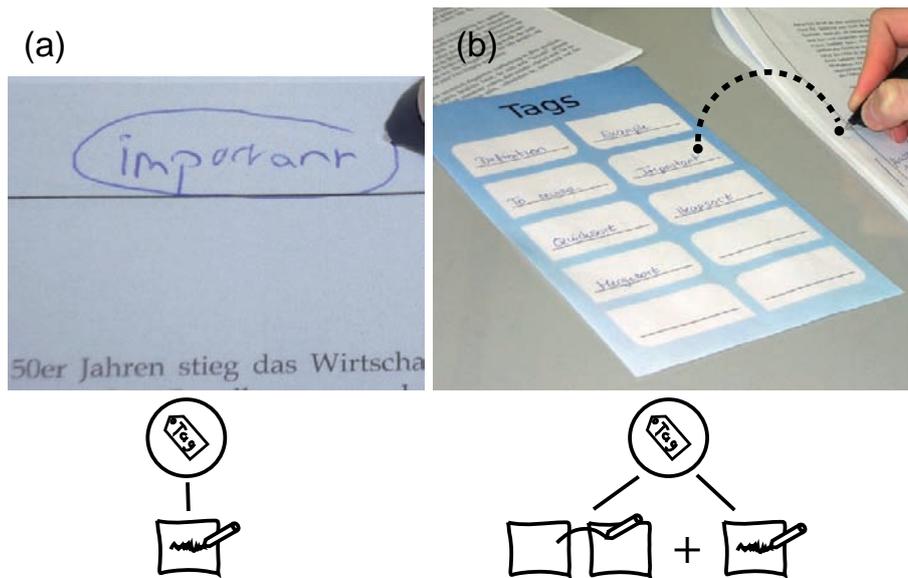


Figure 5.34: Two interactions for applying tags that have been defined on a Tag Menu Card.

document is defined the same way as when creating hyperlinks (see Section 5.4.2). Depending on the type of association area where the tag is written on or associated with and depending on the optional passage marking, the tag applies to a different scope. As for hyperlinks, this interaction technique supports very flexible scopes. It can be used for tagging entire documents, passages within documents as well as collections of documents.

Any time, the user may print a new version of a Tag Menu Card, in which previously defined tags are ordered and sized according to their frequency (tag cloud). Optionally, tags defined by all users or by members of the own learning group can be included.

Collecting all tags on a separate Tag Menu Card has the advantage that the user can immediately access a set of all her tags. In addition, Tag Menu Cards support operations on the tag set (renaming etc.) which can then be automatically applied to the electronic representations of all documents and their subsequent printouts. Moreover, the approach supports co-located collaboration by allowing users physically sharing cards. Finally, research shows that a key factor for the convergence of tags is that the system suggests frequent labels [GH06]. Yet, computer support cannot be assumed in a paper-only environment without a nearby display. In such a context, the Tag Menu Card fosters similar effects as the suggestion of frequently used tags: users will be inclined to re-using tags already entered on the card whenever possible, since the effort is lower than making a new tag entry.

## Visualization

Own and shared tags are displayed in the viewers for individual documents similarly to hyperlinks. Moreover, all tag concepts are automatically included in the collaborative context visualization as oval nodes (see Fig. 5.9 on p. 125). All documents, passages and collections of documents that are tagged with this concept are connected to it by an edge. This enables the user to quickly get an overview on all these contents.

### 5.5.3 Tagging with Buttons

In Section 5.3.3, we have described how buttons from a toolbar, which is printed on documents, can be used for defining the visibility of an annotation. CoScribe utilizes the same interaction technique and visualization for enabling learners to tag individual annotations with semantic categories. The interaction technique is illustrated by Fig. 5.16 (p. 138) and Fig. 5.17 (p. 139). This interaction is similar to applying tags using Tag Menu Cards with the difference that categories are predefined and that the buttons are printed on each document page allowing a quicker access.

CoScribe currently supports four semantic types, which were derived from the needs identified in our field studies. These are *Important*, *Question*, *To do* and *Correction*. Our approach also allows the user to select own categories to be contained in the toolbar. However, if the user introduces new types, sharing and aggregation of annotations is more complicated due to the larger number of classes which moreover other users might not understand.

### 5.5.4 Tangible Tagging of Processes

All tagging strategies presented so far are used for tagging *documents*. In order to structure and augment the collaborative work *process*, we propose a new concept: tagging processes using tangible objects. This supports users in jointly tagging the temporal phases of a co-located meeting. Such meetings often follow rather implicit, but predefined phases, for example starting by collecting all topics to discuss in a meeting, then discussing each topic in turn and finally planning tasks to perform until the next meeting.

Enabling the user to make these implicit phases more explicit by tagging processes has two main advantages. First, the system can temporally index and structure the users' activities (use of documents, annotations, links and tags) by attributing them to individual phases. This enables a quick access which is more systematic than a temporal access which merely uses an unsystematized timeline. Second, this tagging acts as a meta-cognitive learning scaffold. Unexperienced learners might not know how to efficiently structure a meeting. To state two examples, they might not define clear goals for the meeting or they might be unaware of the importance of noting main results. Offering specific types of phases which have proven to be

efficient can stimulate users to follow these phases that are important for a successful meeting. Stimulating users to explicitly specify the phase they are currently in moreover supports the meta-processes of negotiating and making joint decisions about the process structure.

### Interaction

The interaction technique aims at enabling users to specify the type of the process phase in which they are at the moment. Each time, the group proceeds to a subsequent phase, the users indicate the new phase. For this reason, the system can temporally structure the users' activities (use of documents, annotations, links and tags) by attributing them to individual phases.

Main requirements for the interaction technique are first that it gives equal weight to all participants and enables all of them to specify types. This stimulates negotiation and joint decisions of the group. Second the technique shall produce only minimal extraneous cognitive load in order not to disrupt the actual meeting. Moreover the currently selected phase shall be clearly visible to all users with the goal to prevent wandering from the subject or even reverting to an unstructured meeting.

These requirements are met by a shared tangible object which can have different physical states and is located between the users on the table. A tangible is easy and intuitive to manipulate and its physical state indicates the current phase. We distinguish tangibles which stimulate by their affordances to follow a fix sequential structure from tangibles which suggest phases but do not suggest a specific sequence of them.

In order to illustrate the concept of tangible structuring of phases, we have developed two first prototypes of tangibles and an exemplary process model for learning group meetings.

**Process cube.** The first prototype is a tangible cube. of about 10 cm edge length, which is entirely covered with paper containing the Anoto pattern (Fig. 5.35). On each side of the cube, the paper covering has a different color, representing two up to six different phases in the collaborative process. Each time the group proceeds to a subsequent phase, the cube can be rotated to have the corresponding side lying on top. The cube does not suggest a specific sequential structure of phases, since it can be arbitrarily rotated.

As a shared object of all learners, the cube can be manipulated by all participants. The tangible three-dimensional cube provides for a very intuitive, shared representation and gives equal weight to all viewpoints of the participants located around it.

The system detects the cube's current state by sensing RFID tags which are located inside it. This automatically creates an index on a timeline in the context view, enabling the user to get an overview on the temporal structure of the meeting and to easily access all contents and activities of specific phases.



Figure 5.35: The Process Cube for collaborative tangible tagging of group processes

Documents can be linked to specific phases performing the hyperlink gesture on the corresponding side of the cube. Moreover, Digital Paper Post-its can be attached to the sides. This supports collaboratively collecting semantic items that are important and specific to a phase, e.g. topics to discuss in that meeting.

Based on existing models for structuring collaborative learning [Kin07], we developed the following exemplary model that distinguishes six semantic phases of the learning process (Fig. 5.36):

1. *Startup* (red): participants gather together and identify the topics of the meeting; Post-its represent topics.
2. *Explaining* (yellow): for the basic understanding of a topic, definitions and explanations are discussed; Post-its represent important properties of the topic.
3. *Elaboration* (orange): recall cues are collected (e.g. additional details, generated examples and images); Post-its represent recall cues for the topic.
4. *Abstraction* (gray): a topic is put in relation to existing knowledge and the entire learning content; Post-its represent links and levels of abstraction.
5. *Successfully accomplished* (green): contains the topics that have been successfully discussed and accomplished; Post-its represent topics.
6. *Ending/To Do* (blue): end phase of the meeting; agreement on tasks to perform individually after the meeting; Post-its represent to do items.



Figure 5.36: Exemplary model of semantic phases for structuring learning group meetings

Depending on the group processes to support, the cube can be easily modified to support other semantic types.

**Process knob.** In contrast to the process cube, the second prototype illustrates a tangible object that stimulates users to follow phases in a specific sequence. It is a rotary knob that can be rotated by 360 degrees. Around the knob, fields for the phases are arranged. Depending on the rotation, the needle points to one specific field indicating one phase. The rotary knob is depicted in Fig. 5.37.

Similarly to the cube, this tangible provides for intuitive manipulation of a clearly visible, shared representation. The circular arrangements affords rotating the knob



Figure 5.37: The Process Knob for collaborative tangible tagging of group processes

|                                       | Digital Paper<br>Bookmarks | Tag Menu<br>Card           | Tag Buttons | Process<br>Tagging |
|---------------------------------------|----------------------------|----------------------------|-------------|--------------------|
| Documents / Process                   | D                          | D                          | D           | P                  |
| Free Tagging                          | •                          | •                          | -           | -                  |
| Category Tagging                      | •                          | -                          | •           | •                  |
| Paper Documents                       | •                          | •                          | •           |                    |
| Digital Documents                     | -                          | •                          | •           |                    |
| Scope                                 | Page                       | Document,<br>Page, Passage | Annotation  | Phase              |
| Full visual feedback without computer | •                          | -                          | -           | •                  |

Table 5.7: Comparison of interaction techniques for tagging

a little further to make the needle point to the adjacent field each time the users proceed to the subsequent phase. Documents can be linked to specific phases as well, since the fields act as association areas for hyperlink gestures.

In this section, we have introduced a novel approach for collaboratively tagging the process structure in co-located meetings. Subsequent research in the fields of design and computer science should identify further types of tangibles that meet the requirements. Research in social sciences should develop further process models and validate them.

### 5.5.5 Comparison

We have presented four different interaction techniques for tagging documents and processes. These techniques complement each other, serving different purposes and having different properties. The choice of an appropriate technique depends both on the task and the user's personal preferences. Table 5.7 gives a comparative overview of these techniques.

## 5.6 Conclusions

This chapter presented CoScribe, a concept and system for collaborative knowledge work. We introduced a coherent set of novel pen-based and tangible interaction techniques that, by their interplay, provide sophisticated support for the key elements of collaborative knowledge work identified in Chapter 2. These consist of taking handwritten notes, of making handwritten annotations on documents, of creating and following cross-media hyperlinks between documents and of tagging documents and processes. Unlike most other technologies for knowledge work, CoScribe supports not only digital documents but enables the user also to work with

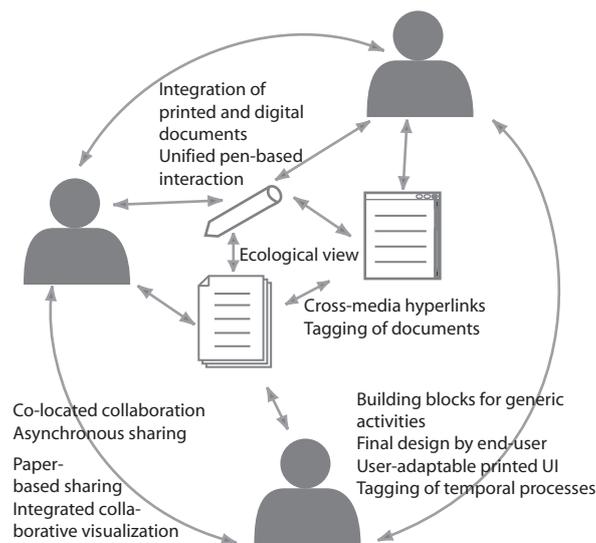


Figure 5.38: Mapping of CoScribe's main concepts to the ecological perspective

printed documents. Interactions on printed documents are automatically digitized. This turns paper into a digitally augmented interactive surface and provides for combining the intuitive and flexible interaction on paper with digital support. Several users can work with CoScribe in a co-located setting and users can share their user generated content (made on paper or on digital documents) with co-workers over distance. The design of CoScribe was guided by an ecological perspective on knowledge work and by our theoretical model of interaction with Pen-and-Paper User Interfaces.

**Ecological perspective on knowledge work.** This holistic socio-technical approach directed our attention to the various relations that constitute an information ecology. These are relations between multiple users, relations between a user and the documents and tools he or she interacts with, relations between multiple documents and relations between the user and the practices of a given work setting. We now summarize by which means CoScribe supports each type of relation. These means are depicted in Fig. 5.38.

*Relations between multiple users.* CoScribe puts an emphasis on collaborative settings. Users can both engage in co-located collaboration and share documents and user-created content over distance. Collaborative content can be accessed in the digital viewer or can be printed on paper. To the best of our knowledge, CoScribe is the first contribution that comprehensively addresses the challenges that are related to asynchronous sharing of handwritten content on paper. This comprises both the processes of sharing own annotations and of accessing shared annotations of other

users. We presented a mechanism that enables users to share annotations directly on paper. This is seamlessly integrated with paper-based annotation. Moreover, we examined how shared handwritten annotations of many users can be visualized in an integrated manner and proposed a novel visualization that leverages flexible zooming for integrating annotations of all users.

*Relations between the user and documents/tools.* CoScribe unifies the interaction with physical and digital artifacts. Regardless if a document or a tool is available in printed form or visualized on a display, the same pen can be used to interact with it. As a consequence, the user does not need to switch between different input devices. One can even perform interactions that span paper and displays, for example linking a Web page with a printed document with one single pen gesture. This tight coupling stands in contrast to most other paper-based applications and improves over augmented desk systems, which restrict the interaction to a small digitally augmented region.

*Relations between multiple documents.* Through its ecological perspective, CoScribe focuses on the relations that exist between multiple documents of a document collection. We presented novel interaction techniques that allow to express these relations in order to integrate and structure hybrid collections of printed and digital documents. A first technique provides for creating, following and sharing hyperlinks between any combination of printed and digital documents. In contrast to previous work on cross-media hyperlinks, our technique supports the wide range of document types which is typical for information ecologies. This includes various formats of digital documents, their printed versions, Web pages and physical books. Moreover, we contributed three novel techniques that offer support for structuring document collections with both predefined and freely-chosen tags. A fourth technique tackles the question of how temporal processes can be tagged by collaborating users with tangible objects. We opened this research direction and proposed two first examples of interaction techniques. Finally, we presented a collaborative ecological view that visualizes the relations within the entire document collection. It provides a high-level overview on all hyperlinks and tags and a structured access to the documents of the collection.

*Relations between the user and work practices.* The theory of Information Ecologies postulates that technology should be designed in a way that leaves freedom of design to the end user. This allows end users to adapt technology in order to symbiotically integrate it into the given practices of a local information ecology. In response to this postulation, CoScribe supports the generic activities of annotating, linking and tagging. As we have discussed, these can be used in very flexible ways for various purposes in different settings of knowledge work. As a consequence, CoScribe is not tailored to a specific purpose but offers a set of generic tools that the end-user can flexibly combine and use according to his or her needs. This aspect moreover directed our attention to the paper user interface, which in previous work is static and cannot be adapted by the end-user. In contrast, CoScribe offers a user-adaptable paper interface.

**Theoretical interaction model.** Using our interaction model of Pen-and-Paper User Interfaces as a basis for the design of CoScribe's interaction techniques prove beneficial. It lead to simple and reliable interactions, to a varied and rich user experience as well as to a high degree of direct manipulation.

*Simple and reliable interactions.* The interaction techniques are inspired by traditional practices of working with documents and let the user maintain her existing practices of working with a pen and printed documents. For instance, users can make handwritten annotations at any position within documents and create handwritten references and tags. This accounts for the highly individual practices observed in contextual inquiries. A small set of more formal interactions can be used to inform the system about the semantics of these informal artifacts. These interactions are based on a small set of recurrent core interactions. This leads to interaction that are easy to learn and to use and that are reliable even with the restricted feedback capabilities of Pen-and-Paper User Interfaces.

*Rich user experience.* We have shown in this chapter that the novel interaction techniques draw upon the large variety of traditional paper-based practices. This includes writing with a pen on paper, manipulating the orientation of physical objects and creating physical arrangements of multiple documents, e.g. in paper folders. Further techniques leverage tangible interaction by attaching physical stickers on a paper document and making pen-based associations between different documents. These varied interactions stand in contrast to paper interfaces that mimic the interaction of Graphical User Interfaces and are restricted to the interactions of writing, drawing symbols and clicking on single paper surfaces.

*High degree of direct manipulation.* The user can directly interact with documents using a digital pen and – for physical documents – his or her hands. This stands in contrast to the indirection caused by using a mouse. The interaction techniques leverage different types of paper-based tools, e.g. for defining collections of documents, for indexing document pages and for tagging documents. All these tools are not only instruments but also objects of interest, as they contain and represent first-class information. In contrast to typical tools in GUIs such as menus or toolbars, this double character lets the user perceive working directly on and with first class objects instead of interacting with tools. This results in a high degree of direct manipulation.

The CoScribe concept has been implemented in a working prototype and evaluated in three user studies. The next chapter will discuss implementation issues before we will present the results of the user studies in Chapter 7.



# 6

## Implementation

This chapter provides a brief overview on the technical realization of a working prototype of CoScribe. Our prototype is a client/server system that includes both software components and specific hardware components. Figure 6.1 gives an overview on the components. With a few exceptions, which are mentioned below, CoScribe is implemented in Java. Most components of CoScribe are executed on a local client. A database that is deployed both locally and on a central server is used for storing data and sharing data with other users. In the following, we will briefly describe the components.

### 6.1 Document Representations

#### 6.1.1 Printed Documents

**Print toolkit.** The print toolkit is a software component that allows to create printouts of documents and printed tools. Users can print a document with one of various layouts. These define for instance how many document pages and empty notetaking areas are included on one printed sheet of paper and at what places these are positioned.

This relies on the concept of logical and physical pages. A logical page corresponds to one page of a physical document. Several logical pages can be combined on one physical sheet of paper. Each logical page can be freely scaled and positioned on the physical page (see Fig. 6.2). Moreover, within each logical page, the actual page of the digital document can be freely positioned and scaled. This provides for leaving empty areas for notetaking besides, above, below or around the document page.

The toolkit adds the Anoto pattern to the printouts and automatically manages the association between the digital contents and their physical positions on the

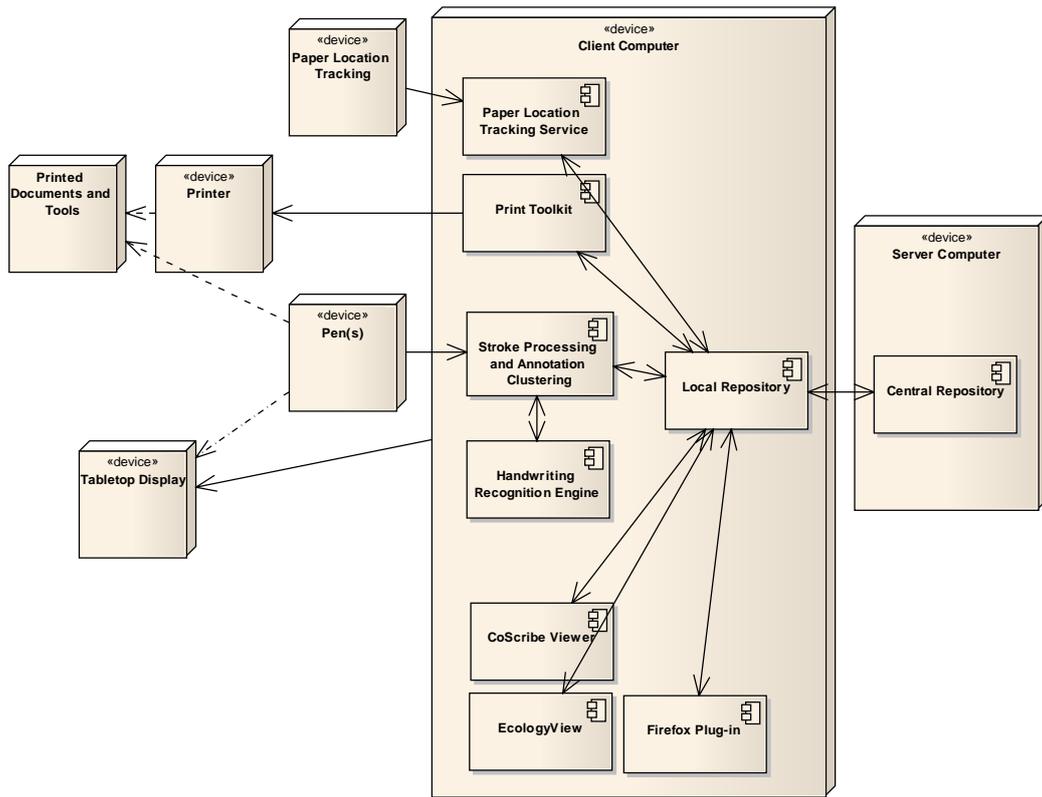


Figure 6.1: Overview on the software and hardware components of CoScribe

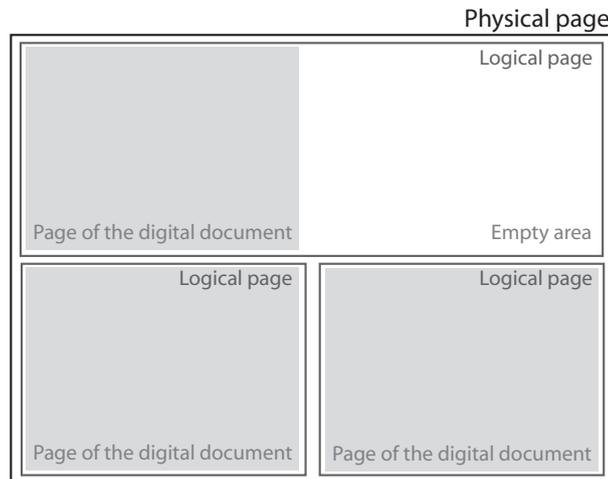


Figure 6.2: Example of physical and logical pages

printouts. This is necessary for mapping pen interactions on paper to the underlying digital contents. The mapping information is stored in the repository.

**Printer, printed documents and printed tools.** The Anoto pattern can be printed on standard office paper along with the document contents in one single step using commercially available standard printers. Empty pages which contain nothing but the Anoto pattern can be printed with monochrome printers. However, if the pattern is to be printed on top of other document contents, a color printer with four-color separation is required for enabling the pen to distinguish the pattern from the document contents. In this case, the *pattern* is printed with black toner, while black *document contents* are printed by combining cyan, magenta and yellow toner. Black toner contains carbon that absorbs the infrared light emitted by the LED of the Anoto pen and makes the pattern dots appear black for the camera. In contrast, cyan, magenta and yellow do not absorb infrared light to the same degree and appear white for the camera. For printing documents, we use an OKI C5900 [OKI] color laser printer.

Currently supported document types for printing are PDF documents and PowerPoint slides. Tangible tools that are printed with the toolkit (or could be offered for purchase by paper suppliers in the future) comprise Digital Paper Bookmarks, Tag Menu Cards, folders and empty sheets for notetaking.

### 6.1.2 Digital Document Viewers

CoScribe includes digital viewers for documents and their relations. This comprises the CoScribe viewer for digital versions of PDF and PowerPoint documents, a plug-in for Mozilla Firefox for hyperlinks and tags on Web pages and an Ecological view that enables an overview on all documents, links, tags and users.

The CoScribe viewer and the ecological view are implemented in Java.

The Firefox plug-in is realized in JavaScript. The script is executed by Greasemonkey [Gre] on each Web page which is accessed in Mozilla Firefox. The underlying principle is that our plug-in is lightweight and manages only the 'view' and 'control' elements of the Model-View-Controller paradigm. All information is managed by CoScribe. The plug-in therefore communicates with a CoScribe Firefox Service (which is implemented in Java and has access to the repository). The communication is performed over the MundoCore middleware [AKM07] using a MundoCore plug-in for Firefox [SHM09], which was developed at our institution.

The plug-in displays the Web page in an IFrame, maintaining the normal visualization of and interaction with Web pages. In addition, the plug-in visualizes association areas for CoScribe hyperlinks and tags and visualizes existing links and tags. CoScribe tracks the pen interaction on these association areas. If a new link or tag is created or a link hot-spot is tapped on, the CoScribe Firefox service informs the plug-in about new contents or requests to display a new Web page.



Figure 6.3: Two prototypes of pen-enabled tabletop displays

## 6.2 Pen-enabled tabletop display.

CoScribe comprises a tabletop display which supports input with the same Anoto pens as used on paper. We constructed two display prototypes. A first prototype (Fig. 6.3 left) is particularly suited for individual work. It is of medium size with a diagonal of 82 cm and is slightly inclined for improving readability. A second prototype (Fig. 6.3 right) addresses co-located collaboration by multiple users. The diagonal of this display measures 112 cm. It is entirely horizontal to enable interaction from each side.

The technical setup of the prototypes is illustrated in Fig. 6.4 and in Fig. 6.5. The display uses back-projection with a full HD resolution beamer. During our experiments, we were using an Optoma HD 80 beamer. To enable pen interaction, the image is projected onto a specific foil which the Anoto pattern is printed onto. Following the approach of [BHH<sup>+</sup>07], we used a HP Colorluculent Backlit UV foil and printed the Anoto pattern onto it. This foil is put between a supporting plexiglass layer of 5 mm width and a layer of 1 mm width, which protects the surface. This approach is cheap and applicable to displays of any size – ranging from small to wall-size displays.

Since we reserve a particular region of the pattern space for the display, the system can distinguish between pen interaction on paper and pen interaction on the display. The Anoto coordinates of pen events originating from the display are automatically converted to screen coordinates. If the event is on a window of CoScribe, the system directly processes the event. Otherwise, a standard mouse event is generated. This enables controlling third-party applications with the pen. Several pens can be used simultaneously on the same display.

A limitation of the current prototypes is that the Anoto pens leave slightly visible ink marks on the protecting layer, which however can be easily wiped off. A next version of the prototype should include an ink-repellent surface. Alternatively, future models of Anoto pens could control the ink cartridge and emit ink only if used



Figure 6.4: Setup of the display prototype

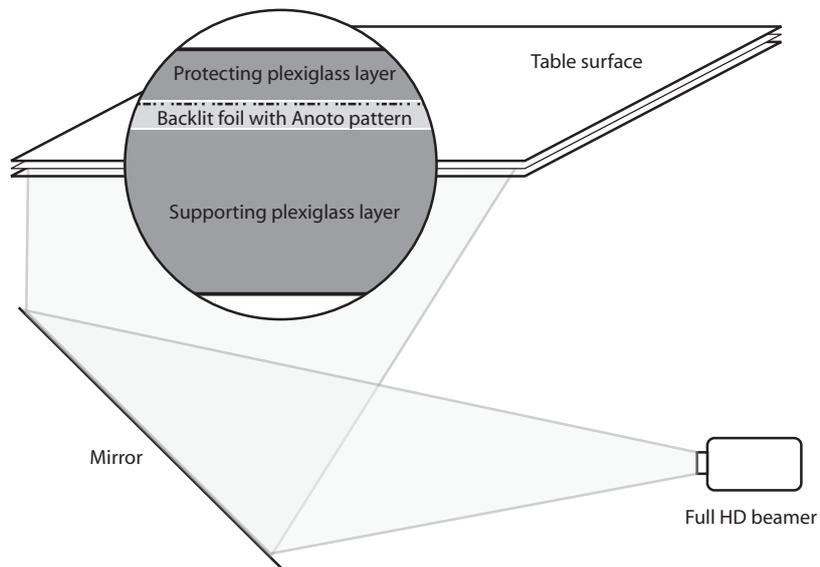


Figure 6.5: Schematic illustration of the technical setup

on paper.

### 6.3 Physical Interaction Tracking

**Pen(s).** Our prototype uses Anoto pens of the models Logitech io2 [Log] and Nokia SU-1B [Nok]. These pens are used to interact on printed documents and tools as well as on a tabletop display. Pen data is transferred to the client computer using one of the following two ways: Either data is temporarily buffered on the pen until it is put into a reading device that transfers all data to the client computer using a USB connection. Or data is streamed in real-time to the client computer using a Bluetooth connection with the SPP profile. When the pen is utilized on a display, real-time streaming is used. In the Bluetooth setup, up to 8 pens can simultaneously connect to one single client computer. This provides for co-located use. If a larger number of pens is to be used, several client computers can be used in the same place, which synchronize their data via the central database.

**Paper location tracking.** CoScribe optionally uses camera-based tracking of the locations of paper documents and folders. For this purpose, we use ARTag [Fia04] which tracks locations using fiducial markers that are printed on documents and folders.

### 6.4 Pen Data Processing

**Stroke processing.** Once the pen data is sent to the client computer, it is automatically handled by the stroke processing module. The raw pen data contains for each pen stroke a set of sample. Each sample is a tuple consisting of the page id, the (x,y)-coordinate on that page, the pen tip force and the timestamp. The stroke processing module checks if the strokes originate from paper or from the display and attributes it to the associated digital contents or interface elements. It then interprets the interaction and executes a command, if necessary.

If the stroke is not a command but an inking interaction, individual strokes are clustered into annotations. An annotation is a set of strokes which are temporally and spatially contiguous. CoScribe comprises a component for clustering that uses an on-line algorithm for both spatial and temporal clustering. This is presented in more detail in the next section.

In contrast to other applications, which usually model handwritten annotations as being located on one single page, our model supports annotations that span multiple pages. This is particularly important, since one physical page can contain multiple logical pages and users frequently write over the boundaries of individual logical pages. It would not be acceptable to cut these annotations into two or more separate fragments. Instead, our model keeps these annotations intact and treats them as one single annotation. The x and y coordinates of the individual samples of the annotations are normalized to be in a [0..1] range, where the coordinate (0,0)

denotes the upper left and  $(1,1)$  the lower right corner of the bounding box of the annotation. This makes the modeling independent of the actual position of the annotation on a page. For each logical page the annotation is located on, it contains some context information that indicates which portion of the annotation is visible at what area of this page. This makes the annotation model very flexible.

**Handwriting recognition.** New or updated annotations are sent to the handwriting recognition engine which is part of Microsoft XP Tablet PC edition. This is encapsulated as a separate server component that can be executed either on the local computer or – if an operating system other than Windows is used – on the server computer. The recognized text and up to 10 recognition alternates are stored in the repository along with the annotation.

## 6.5 Annotation Clustering

### 6.5.1 Online Algorithm

CoScribe uses a two-stage online algorithm for the temporal and spatial clustering of strokes into annotations. First, two consecutive strokes are clustered into the same annotation if the temporal gap between the ending of the first and the beginning of the second stroke is smaller than 3500 ms and if both strokes are located on the same page. Second, the spatial clustering proceeds as follows. It is detailed in pseudocode in Algorithm 1. The new stroke is compared to all other annotations that are candidates for a merging action. Typically, candidate annotations are all annotations on the same physical page, but it is also possible to take into account annotations located on other pages. This provides for annotations that span multiple pages. For each sample of the new stroke and each other annotation, the euclidian distance to the nearest sample of the other annotation is calculated. Next, the mean of all these distances per annotation is calculated, whereby only samples are considered that are located at the side of the annotation that is oriented towards the other annotation. This restriction is realized with a threshold value for distances. If this nearest mean distance is below a maximum-distance threshold, both annotations are clustered into one annotation.

The worst case complexity of this algorithm is  $O(st \cdot s \cdot max_s)$ , here  $st$  denotes the number of new strokes,  $s$  the total number of samples contained in all strokes on a page and  $max_s$  the maximum number of samples contained in one single stroke. We opted for this more complex approach and against using bounding boxes because the rectangular bounding box is often very different from the actual shape of an annotation (e.g. large circles or diagonal lines and arrows). The latter approach would merge many annotations which are not related to each other.

---

**Algorithm 1** The spatial clustering algorithm used in the working prototype of Co-Scribe

---

**input:** a set of new strokes, each consisting of a set of samples  
**input:** a set of annotations, each containing a set of strokes each containing a set of samples

```

for each new stroke  $st$  do
  for each sample  $s_{st}$  of stroke  $st$  do
    for each annotation  $a$  do
      for each sample  $s_a$  of annotation do
        calculate the euclidian distance  $d_{s(a,st)}$  between  $s_a$  and  $s_{st}$ 
      end for
      store the nearest distance as the minimal distance  $d_{a,s}$  between the sample  $s_{st}$  and annotation  $a$ 
    end for
  end for
  for each annotation  $a$  do
    calculate the mean  $\bar{d}_a$  of all distances  $d_{a,s} < d_{thr}$  for all samples  $s_{st}$  of stroke  $st$ 
  end for
  select annotation  $a_{nearest}$  with the lowest value  $\bar{d}_a$ 
  if  $\bar{d}_a < d_{max}$  then
    add stroke  $st$  to annotation  $a_{nearest}$ 
  else
    add stroke  $st$  as a new annotation  $a_{st}$ 
  end if
end for

```

---

### 6.5.2 Evaluation

**Method.** We evaluated the performance of the algorithm. During three regular university lectures, we collected annotations on lecture slides made by 29 students (see Section 7.2 for more details on the method). These were automatically clustered using our clustering algorithm. From the resulting set of 1983 annotations, we randomly selected a subset of 150 annotations for this evaluation.

The clustering of annotations is a complex task which cannot be made unambiguously. It may be hard even for human raters to decide if the strokes are correctly clustered or not. For example, one person might consider a set of strokes as two independent annotations (e.g. a sketch and some text) while another person might consider this as one single annotation. For this reason, we decided to use an inter-rater agreement method. This measures the agreement of two human raters and allows to assess if the human classification is reliable. Two independent human raters classified each annotation. Each rater classified the annotation as incorrect if in her opinion it contained one or more strokes of another proximate annotation or if a proximate annotation contained one or more strokes of this annotation. Otherwise, the annotation was classified as correct. In order to gain more agreement

between both raters than by following merely subjective classification criteria, the raters used the following rough guidelines:

- An annotation must have a coherent content on a semantic level.
- The strokes forming one annotation must be spatially coherent, i.e. two clusters of strokes that have a large distance are considered as two separate annotations.
- Clusters of strokes that are visually separated in an explicit manner (e.g. by a line drawn by the participant) are considered as separated annotations.
- Sketches are own annotations. However, if text can be attributed unambiguously to a sketch, both this text and the sketch is considered as one annotation.
- An annotation has always one type and one visibility level. Strokes of different types and visibilities belong to different annotations.

In order to measure the inter-rater agreement, we used Cohen's kappa [Coh60] measure.

**Results and discussion.** Figure 6.6 depicts the performance of the algorithm as judged by two independent raters. A total of 60 % of the annotations were classified as correct by both raters. 24.7 % were classified as incorrect. The remaining 15.3 % of the annotations were rated differently by both raters. The score of Cohen's Kappa was 0.68. This indicates high inter-rater reliability.

The proportion of one quarter of annotations that was clustered incorrectly by the algorithm might seem high. However, an annotation was classified as incorrect

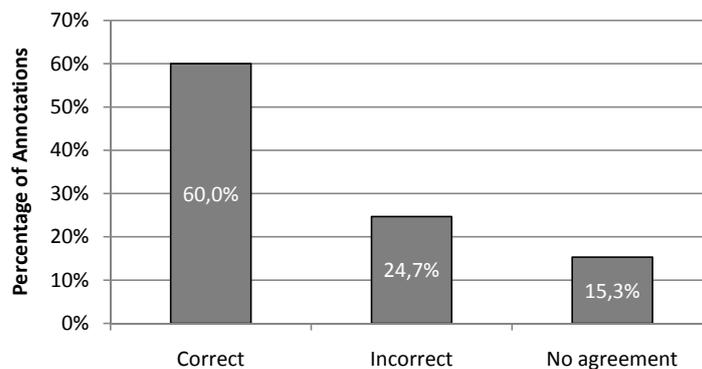


Figure 6.6: Percentage of annotations rated as correctly or incorrectly classified by both human raters and of annotations on which the two raters did not agree ( $N = 150$ )

even if only one single stroke (e.g. a separation line or the dot at the end of a phrase) was separated from or added to the annotation. In practice, these minor errors are not problematic. Moreover, a manual analysis of error cases showed that the human raters frequently classified clusters as incorrect that contained strokes that were spatially distant although being contiguous on a semantic level. This was a problem of the classification guidelines.

There remains nevertheless a substantial proportion of annotations that are not correctly clustered. Moreover, the percentage of 15.3 % of annotations on which the raters disagreed shows that a correct clustering is a hard problem *per se*. For this reason, it is very important that the user has the possibility to manually correct the automatic clustering. A second implication is that critical functionality should not rely on automatic annotation clustering where no feedback on the clustering is provided to the user. Let us consider the example of a user who would classify a set of strokes – which she considers as being one annotation – as private. But the algorithm would cluster these strokes into two annotations. Hence, the system would treat one of them as private, the other would have default visibility. Such a system behaviour is clearly not acceptable. In this case, the user should have the possibility to manually define the extent of the classification on paper, for instance by encircling the area. On the contrary, clustering can be efficiently used for less critical functionality like the semantic classification of annotations. In this situation, the fact that a classification applies to a smaller or larger set of strokes than expected by the user is not critical. Instead of an entirely predictable system behavior, the user might prefer an interaction that is quick and easy because it is not necessary to manually define an area the classification applies to.

Finally, it must be stated that our goal was to find an on-line algorithm which performs reasonably well for a working prototype but not to make a scientific contribution to the field of handwriting or sketch segmentation. It is certain that, with more effort invested, more powerful algorithms can be found.

## 6.6 Data Persistence

All data is stored in a graph-based database system, which was developed at our institution. This system comprises a central repository on a server, which is accessed by all local installations of CoScribe. Moreover, a local mirror of this repository (containing of all information the specific local user has the right to access) is deployed on each client computer. The database system automatically manages the transfer of information between the local mirror and the central repository. If no connection can be established, new or updated data is synchronized once the connection is available again. This provides for using CoScribe in a single-user configuration without central server. Moreover, CoScribe can be used even when the central server cannot be accessed, for example when the user is not in her office and has no internet connection.

# 7

## User Studies

This chapter presents and discusses the results of three user studies of CoScribe. These aim at evaluating CoScribe by verifying if our contributions meet the requirements identified in Chapter 2. Each of the three user studies focuses on a specific setting of learning at universities. In Section 7.2, we present a study that assessed the use of the printed user interfaces for annotating lecture slides. A total of 29 university students participated to this study and used CoScribe during one of their regular lectures. In Section 7.3, the results of a lab study are discussed which focused on the process of reviewing lecture notes after class. This includes the combined use of printed lecture slides and of their digital counterparts as well as of the multi-user visualization of shared annotations. Finally, in Section 7.4, we present the results of a user study that assessed the use of CoScribe for integrating information from hybrid collections of printed and digital documents, as it is for example necessary for preparing articles or term papers.

### 7.1 General Design

**Evaluation of the requirements.** Table 7.1 gives an overview in the requirements which we have identified in Chapter 2. All of them are fulfilled by the design. Most of these requirements are fulfilled by the design of CoScribe. However, several requirements cannot be verified in a purely argumentative way. This comprises the requirements R1b (interactions should be reliable even in the absence of digital feedback) and R4a (interactions should produce little extraneous cognitive load). One main goal of the user studies was therefore to examine if the interaction techniques are simple, quick, intuitive, easy to learn and reliable. As shown in the table, we also evaluated other requirements in the user studies to gain insights into how the interaction techniques and visualizations are utilized by users.

| Requirement  | Contribution of this thesis  | Fulfilled by the design | Evaluated with users in study |
|--|--|-------------------------|-------------------------------|
| R1 Support of both printed and digital documents                       | Richer variety of interaction metaphors  | ✓                       | 1,2,3                         |
| R1a Similar interactions in the physical and in the digital worlds     | Unified pen-based interaction  | ✓                       | 3                             |
| R1b Reliable interactions without digital feedback                     | Reliable interactions without the need for handwriting recognition   |                         | 1,2                           |
| R2 Supporting annotating, linking and tagging collections of documents | Richer variety of interaction techniques   | ✓                       | 1,2,3                         |
| R3 Supporting co-located and remote collaboration                      | Concepts for paper-based sharing and integrated multi-user visualizations                                      | ✓                       | 1,2,3                         |
| R4 Symbiotic integration into current information ecologies            | Integrated solution based on ecological perspective  | ✓                       | 1,2,3                         |
| R4a Interactions that produce little extraneous cognitive load         | Easy and intuitive pen gestures from a small inventory   |                         | 1,2,3                         |
| R4b Rich and flexible interactions                                     | Freedom to maintain current practices; adaptable printed user interface; choice amongst interaction techniques | ✓                       | 1                             |
| R4c Structured access to all information                               | Ecological view  | ✓                       |                               |

Table 7.1: Overview of the evaluation of the requirements from Chapter 2

**Evaluation of the user satisfaction and acceptance.** A second main goal of the user studies was to assess user satisfaction in order to examine if the novel interaction techniques and visualizations are accepted by the users. In this respect, subjective feedback was an important instrument. Further, we aimed at gaining first user experiences and feedback on potential for improvements.

The studies deliberately did not assess direct effects that the use of CoScribe might have on learning quality. First, information ecologies comprise manifold sources that influence learning quality. These cannot be entirely captured and can hardly be controlled. Controlled user studies that assess the influence of a few precisely defined independent variables incur the risk of leading to rather unnatural settings. Second, it is hard to define and to measure quality of learning. Many studies measure learning quality by the grade obtained in an exam. However, this leaves aside many aspects which might be equally important for the learning quality, such as long-term effects on the knowledge of the domain, the joy of using a system, the

sensation of rather integrated or rather disruptive workflows, the time needed for carrying out a learning task or the soft skills acquired for example by intense group work.

For these reasons, we decided to ground our user studies on a concept that puts strong emphasis on the fact that the learner is a subject who is free to decide upon her role and about the methods and tools she uses. In contrast to a merely instrumental conception of technology as an effective means for attaining a well-defined learning goal which can be unambiguously measured, this concept perceives technology as a means for opening up novel spaces (in a metaphorical sense). These spaces may enable novel and previously unforeseen ways of learning [Ses04, p. 96 sqq.]. As a consequence, new methods and technologies have first and foremost the character of a proposal which is made to the learner. It is the learning subject who then decides if she accepts this offer. Therefore, it is crucial to assess the acceptance of any new technology before examining its potential influences on the learning process.

## 7.2 Lecture Notetaking

A first user study examined the use of CoScribe for annotating documents in a realistic setting. This setting was annotating lecture slides during regular computer science lectures. Our goal was to assess the ease of learning and the ease of use of the printed user interface and of the interaction techniques for making annotations and for classifying them with visibilities and semantic types. A further question was if the techniques can be easily integrated into the ecology of lecture notetaking that is characterized by a high degree of intrinsic cognitive load. We moreover analyzed the annotations made by the participants in order to assess types and frequencies of annotations made during the lectures.

### 7.2.1 Method

For our study, we selected three computer science courses of different years of study. These were a second year introductory lecture on sorting algorithms, a third year lecture on computational engineering and a lecture for graduate students on ubiquitous computing. A total of 29 students (5 female, 24 male) were recruited among the attendees of these lectures. Their semester of study averaged out at 5.6 ( $SD = 3.1$ ). Participation was voluntary and no compensation was given. To avoid positive bias, none of the students was personally known to us nor did they attend one of our courses.

Before the lecture the participants received an electronic pen and a digital paper printout containing the slides of the current lecture. Each A4 sheet contained two slides on the left-hand side and empty areas for taking longer notes to the right. He or she was trained for three minutes on how to make annotations and to classify them using the digital pen. The task during the lecture was to make annotations on the printout the same way the participants usually do. In addition, we asked them

to semantically classify and/or to share annotations using the paper buttons when it seemed appropriate to them. All these activities were digitally captured by the CoScribe system for later analysis.

After the lecture, the participants filled out a standardized questionnaire for quantitative feedback on the paper-based user interface. The questionnaire contained 25 closed and open questions related to the printed interface, to the digital pen, to the lecture and to personal information. Finally, in each of the three sessions, we conducted a semi-structured group interview with three to six participants. The goal was to gather additional qualitative insights into benefits and shortcomings of the current design and to brainstorm about which further functionality would be helpful. These interviews were videotaped and varied in length from 35 to 70 minutes.

In the statistical analysis of the questionnaire data, we investigated correlations between Likert-scale items which were five-point scaled and performed  $\chi^2$ -tests and t-tests to identify significant group differences. All these tests were based on a level of significance of 95 %. Moreover, we analyzed the annotations that were made during the lecture. This comprised a statistical analysis of frequencies, positions and types of annotations.

### 7.2.2 Results and Discussion

**Document annotation.** Participants considered digital pen and paper for annotating lecture slides to be easy to use (Requirement R1, R4a). All users reported that annotating printed lecture slides with the digital pen worked reliably and as they had expected (R1b). In the questionnaire, the participants judged the use of CoScribe about as distracting as traditional pen and paper ( $M = 2.7$  on a scale ranging from 1=more distracting to 5=less distracting,  $SD = .7$ ,  $N = 29$ ), but significantly less distracting than using a laptop ( $M = 4.5$ ,  $SD = 1.0$ ,  $N = 29$ ) ( $T = 9.0$ ,  $df = 26$ ,  $p < .001$ ) (see Fig. 7.1). This indicates that the novel techniques can be well integrated into the current ecology (R4). Yet, these numbers show not only the ease-of-use of CoScribe's paper interface, but underscore the results of our first field study (see Section 2.3.1). This study found that pen and paper is largely preferred to laptops for taking lecture notes because of the ease of using pen and paper.

Although the users have had only a few minutes for familiarizing with the system and used it during one of their normal lectures, they made a considerable amount of annotations. We collected a total of 1983 handwritten annotations during three lectures. In average, each user made 68 annotations ( $SD = 29$ ,  $N = 29$ ) during a 90 minutes lecture. Figure 7.2 shows a histogram of the number of annotations each user made. A high variability between participants was encountered, ranging from only 7 up to 141 annotations per participant. This is in-line with the findings of our second field study (see Section 2.3.2) and findings reported in the literature [Obe03, Mar98]. However, we found no significant differences between the different lectures. On average each user annotated 63 % of all available slides, making 2.6 annotations on each of these slides ( $SD = 1.7$ ). Figure 7.3 shows a histogram of the average number of annotations that a single user has made on each slide. No

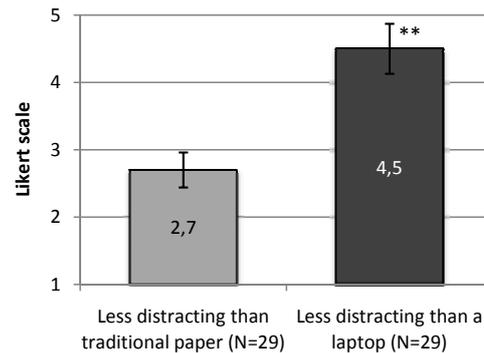


Figure 7.1: Comparing annotating with CoScribe with traditional pen and paper and a laptop. (The error lines indicate the 95 % confidence intervals of the means.)

relation between the time of the lecture and the number of annotations was found. All slides were annotated by at least one participant. The ten most frequently annotated slides contained an average of 41 annotations ( $SD = 5.8$ ), while only 2 of the 122 slides contained one single annotation. This shows that the participants were active throughout the entire lecture and that no temporal “hot-spots” of very high or very low activity existed.

Advantages mentioned in comparison to electronic systems without paper support consisted of the possibility to easily make annotations during class and to work with formulas. For instance one student reported in the interview an experience with a Web site designed for group work of mathematics students, which was not accepted and not used by the students since it was by far too complicated to enter formulae with the keyboard. Advantages with respect to pure paper environments were better orientation in the digital document, quicker access in later semesters without the need to keep a physical folder in range and better support for group work. Several users reported that a drawback of the system is that the Anoto pen has only one fix color. Instead, they would prefer a pen that can switch between various colors. This is an area for improvement of Anoto pens. Another problem was that many users felt irritated because the pen vibrated sometimes when it was used for writing over printed text, which is a technical problem of currently available Anoto pens.

The participants had different preferences concerning the layout of printed slides. This underscores the importance of flexible layouts (R1). Three out of four of the participants reported in the interviews to prefer only two slides per paper sheet, as this leaves free areas for annotations. The remaining participants preferred four or more slides per page. All left-handers wished having the free annotation areas to the left of the printed slides, while right-handers preferred them to the right. Positioning the free areas below instead of besides the slides was judged less appro-

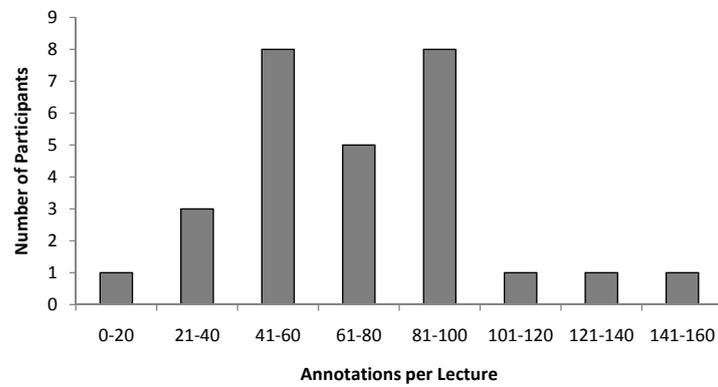


Figure 7.2: Histogram of the total number of annotations made per participant during a 90 min. lecture

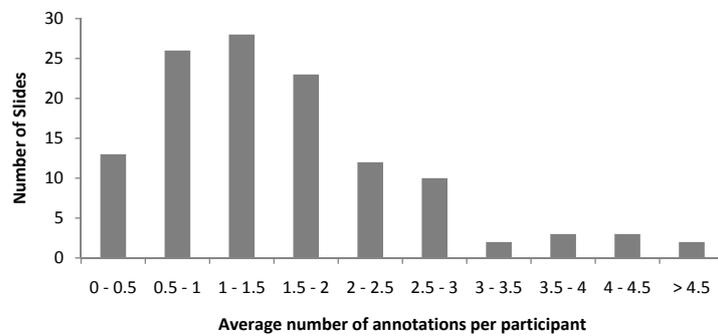


Figure 7.3: Histogram of the average number of annotations each participant made on a single slide

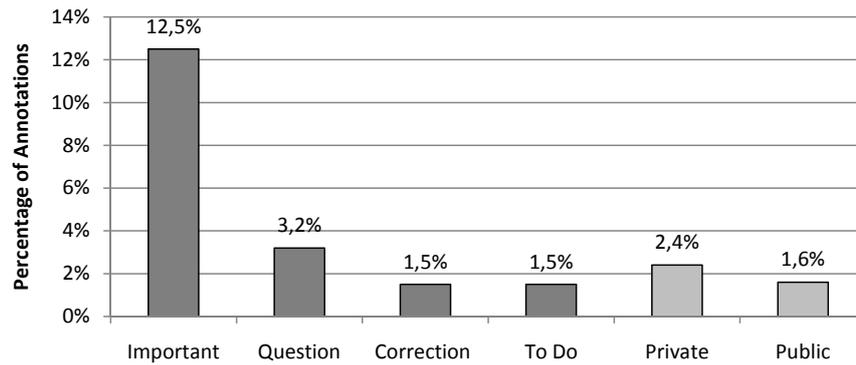


Figure 7.4: Percentage of annotation classified with semantic categories and visibilities ( $N = 1983$ )

appropriate, since the items on the slides are typically organized in vertical order and it would therefore be more complicated to relate notes to individual items.

**Tagging of annotations.** During the lectures, the participants classified a substantial proportion of 18.7 % of the annotations with a semantic category using the Tag Button classification method. This indicates that a few minutes of training is sufficient for learning this technique. The percentage is even higher than the score of 10.8 % of category markers found in our introductory field study of traditional handwritten annotations (see Section 2.3.2). This difference might relate to the fact that in this study, the students used CoScribe for the first time and were motivated to use the new functionality to a higher-than-average degree. Another reason could be that in the study of traditional annotations, we counted only annotations containing a symbol that could be clearly identified as a semantic marker. As there are also other means for conveying semantics, such as keywords or the position where an annotation is made, this score is possibly lower than the score of this study, in which all classifications could be clearly identified as such.

The most frequently chosen category was 'Important' (12.5 %), followed by 'Question' (3.2 %), 'Correction' and 'To Do' (1.5 % each) (Fig. 7.4). Although the categories 'To do' and 'Correction' were used more rarely, several participants of the interviews explicitly stated that these types are nevertheless very important.

Several participants reported the wish to define individual types like 'To revise', 'Definition' or 'Example'. From a designer's perspective, individual types are problematic because types are used to visualize shared annotations. With many different individual types defined by the users, other users would not be able to understand at a glance the meaning of shared annotations. An appropriate trade-off could consist of offering a set of some more common categories, of which the user would choose a subset to be printed on his or her personal printouts. In order respond to

the user feedback and to overcome the limitation of fix categories, we developed the technique of Tag Menu Cards for tagging with freely-chosen keywords and integrated this into the second version of our prototype.

In addition to tagging with categories, we assessed the percentage of annotations that are tagged with visibilities. Tagging with visibilities was performed only for a small percentage of annotations (2.4 % private and 1.6 % public) (see Fig. 7.4). Despite these low scores, we should not conclude that privacy mechanisms are not necessary. In all interviews, it became clear that the system must offer a means for defining the visibility to be of practical use in lecture settings. Rather, the low scores reflect that the default setting of group visibility is appropriate for most annotations.

In the interviews, there was a wide range of responses to the functionality for classifying annotations. While nearly all participants agreed that this is an important feature, that tapping on a button is quick and easy and does not disrupt the main task of annotating (R4), they disagreed about whether the system feedback is sufficient (R1b). Many users reported to feel unsure if a printed button has been correctly activated when tapping on it with the pen. While the pen confirms the pen tap on a button by briefly lighting up a LED, it provides no feedback on the currently activated classification mode. This was not possible with the Anoto pens available at the time of the evaluation. Novel pens with integrated feedback capabilities could make the classification with paper buttons more reliable.

## 7.3 Reviewing Annotations

A second exploratory study assessed the use of CoScribe during review after class. In this setting time is less scarce than during a lecture and learners can make use of the system's entire functionality. We evaluated the use of the CoScribe viewer for collaborative activities, the combined use of paper and digital documents and the appropriateness of Digital Paper Bookmarks for document structuring and cross-media navigation.

### 7.3.1 Method

We recruited nine students (7 male, 2 female) among the participants of the first study. Each participated to a single-user session which lasted about one hour. Participation was voluntary and no compensation was given.

The participant was given an electronic pen, a twenty page printout of slides of an introductory computer science lecture and several Digital Paper Bookmarks. He or she was seated at a table with enough free space for the paper documents. A computer screen on the table as well as a keyboard and a mouse provided access to the CoScribe viewer.

The sessions were structured as follows: At first, the participant was trained for five minutes on how to use the CoScribe viewer and Digital Paper Bookmarks. In the following, we requested the participant to perform given tasks with paper and the CoScribe viewer. This comprised creating annotations and bookmarks on paper

as well as using the software viewer to modify own annotations. Next, we evaluated the appropriateness and compared the multi-user and the single-user visualizations of shared annotations. For this purpose, the participant had the task to browse through shared annotations of other users and to find specific annotations. These shared annotations were made by other participants in the previous study during their computer science lectures. Finally, in order to assess the ability of digital paper bookmarks to support cross-media navigation, we displayed a random slide in the software viewer and asked the student to find the corresponding page in the paper stack using the bookmarks he or she had created before. After the experimental session, feedback was gathered with a questionnaire containing 64 closed and open questions. These covered general questions about the usability of CoScribe, more specific questions on the visualization of shared annotations, on bookmarks, on the paper layout, on the general behaviour in lectures and personal questions. Finally, we conducted a semi-structured interview to gather further qualitative feedback. The entire session was videotaped.

### 7.3.2 Results and Discussion

**Reviewing shared annotations.** Participants reported that to date, handwritten annotations are typically not shared with other students due to the large effort. We asked the participants for what purpose they would use shared notes. Of the variety of answers provided, five users mentioned that they would read the comments made by specific students known to take good notes. Two users stated that notes of different users complement each other, since there is not enough time during a lecture to note all information of importance. Two other users stated to correct own notes with the help of others. In the questionnaire, all participants judged CoScribe to be very helpful for collaboration ( $M = 4.9$  on a 5-point Likert scale,  $SD = .3$ ,  $N = 9$ ).

We evaluated the novel multi-user visualization of handwritten annotations (Requirement R3). For displaying own annotations, the multi-user view is equivalent to the single-user view because the symbols for shared annotations can be easily hidden. Concerning shared annotations of other users, the participants judged the multi-user view as more helpful when seeking an overview of them ( $M = 4.3$ ,  $SD = .9$  vs.  $M = 2.1$ ,  $SD = .9$ ,  $N = 9$ ). This result is highly significant ( $T = -6.64$ ,  $df = 10$ ,  $p < .001$ ). They also judged this view to be slightly more helpful for finding a specific shared comment (Fig. 7.5). We initially feared that the multi-user view would become cluttered, as both own and shared annotations are displayed together. However, participants judged it to be almost as clear as separate views ( $M = 4.3$ ,  $SD = .7$  vs.  $M = 4.6$ ,  $SD = .5$ ,  $N = 9$ ).

To compare the efficiency of both modes, we measured the time needed by the participants to find an arbitrary question and a specific comment on a given slide for each mode. Fig. 7.6 shows that in both tasks, students found them much more quickly using the multi-user view ( $M = 1.1s$ ,  $SD = .4$  vs.  $M = 2.5s$ ,  $SD = 0.5$  and  $M = 4.4s$ ,  $SD = 1.6$  vs.  $M = 8.0s$ ,  $SD = 4.8$ ,  $N = 6$  after having removed obvious

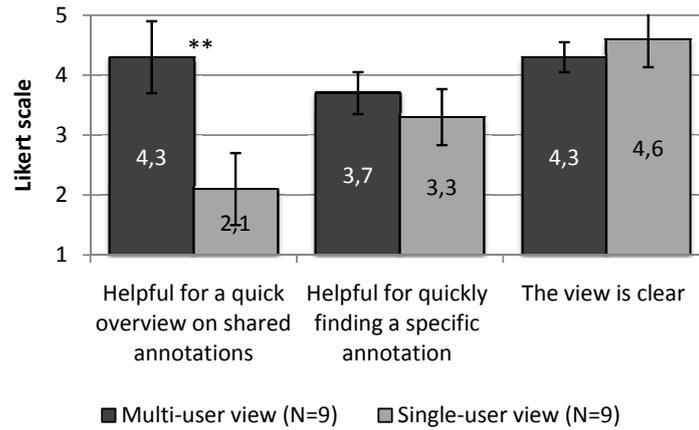


Figure 7.5: Comparing the multi-user view and the single-user view (The error lines indicate the 95 % confidence intervals of the means. \*\* = statistically significant)

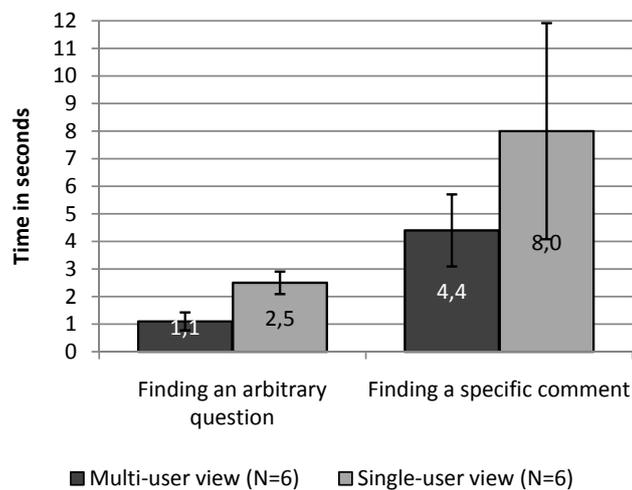


Figure 7.6: Task completion times. (The error lines indicate the 95 % confidence intervals of the means.)

outliers). These data are thus in keeping with the subjective ratings, but can only be considered a qualitative indication due to the small sample size. They support the assumption that the multi-user view enables a fast overview of shared annotations of other users.

Users found that the different icon sizes, with which annotations of other users

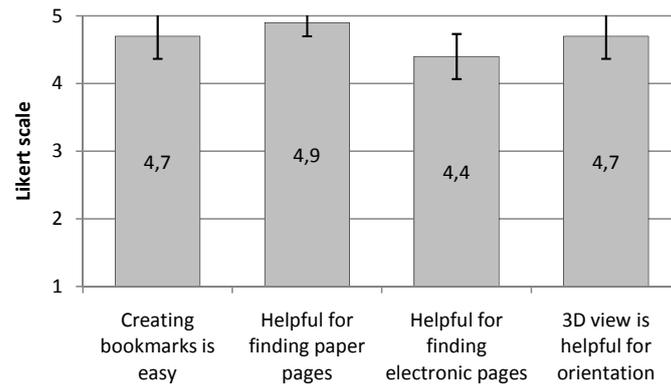


Figure 7.7: Quantitative results for Digital Paper Bookmarks. (The error lines indicate the 95 % confidence intervals of the means.)

are visualized and which reflect the size of the annotation, are helpful, but two users had difficulties recognizing some of the smallest icons and suggested to visualize them in a more highlighted manner. All participants particularly valued the pre-view function, which visualizes all shared comments or all comments of a specific user in a scaled down manner.

In the interviews, three users stated that a list view of all annotations should complement the view to support users in systematically reviewing all annotations. One user suggested a function for filtering out slides that are not important and create a printed summary of the lecture that contains slides that are classified as important.

**Digital Paper Bookmarks.** The observations of how users create and use bookmarks clearly indicate that the interaction technique is highly intuitive and reliable (R1b). All participants readily understood the usage of bookmarks and created and modified bookmarks without assistance (R4, R4a). Seemingly, the bookmarks appeared so professional that two participants asked where they could be bought and if they were expensive. In the questionnaire, bookmarks were rated as being very easy to create ( $M = 4.7$ ,  $SD = .5$ ,  $N = 9$ ) (see Fig. 7.7). Concerning their use, bookmarks were judged to be helpful both for finding a specific page in the paper document ( $M = 4.9$  on a 5-point Likert scale,  $SD = .3$ ,  $N = 9$ ) and for finding a page in the CoScribe viewer ( $M = 4.4$ ,  $SD = .5$ ,  $N = 9$ ). Several participants particularly valued the possibility to fade away from the given structure and to create instead an own structure of the document.

As depicted in Figure 7.7, the visualization of the 3D paper stack with the bookmarks in the software viewer was considered very helpful for orientation within the document ( $M = 4.7$ ,  $SD = .5$ ,  $N = 9$ ). Several users stated that this view provides a better awareness of the approximate current position within the document

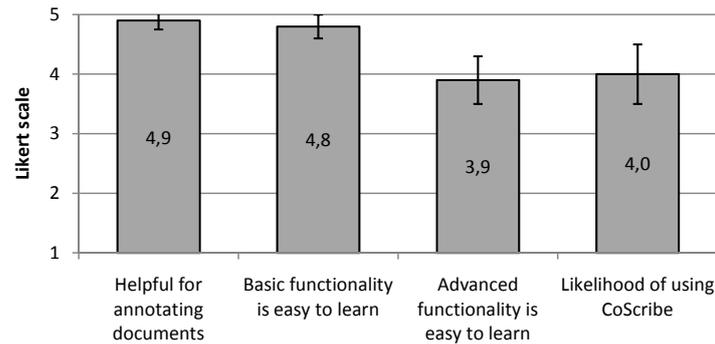


Figure 7.8: General questions on CoScribe. (The error lines indicate the 95 % confidence intervals of the means.)

than the more traditional slide preview scrollpane. Moreover, it takes less space than the scrollpane and provides a representation which closer resembles the real paper stack. However, due to the 3D distortion of the visualization, the text on the bookmarks was less well readable if it was very small or written in a hurry. In this respect, the visualization of bookmarks directly on the slides is more powerful.

These initial findings indicate the efficiency of Digital Paper Bookmarks. They should be confirmed in a subsequent study, in which students work with a larger collection of documents and use various tagging techniques.

**Degree of acceptance.** As discussed above, a main goal of the user studies was to assess the degree to which the participants were willing to accept CoScribe. This indicates if CoScribe has the potential to be used in realistic settings and hence to open up new learning spaces. The positive results of the multi-user visualization in CoScribe and of Digital Paper Bookmarks provides first evidence for acceptance. However, these are rather specific features and their rating possibly cannot be generalized to the entire system. For this reason, we asked some general questions concerning CoScribe.

The participants rated the combination of paper and computing as realized by CoScribe to be very helpful for annotating documents ( $M = 4.9$  on a 5-point Likert scale,  $SD = .3$ ,  $N = 9$ ) (Fig. 7.8). The basic functionality of CoScribe was rated to be very easy to learn ( $M = 4.8$ ,  $SD = .4$ ,  $N = 9$ ). They also judged that the advanced functionality can be easily learned ( $M = 3.9$ ,  $SD = .8$ ,  $N = 9$ ). The overall probability that the participants would use the system once it is available in normal lectures was high ( $M = 4.0$ ,  $SD = 1.0$ ,  $N = 8$ ). These results indicate a high degree of acceptance.

## 7.4 Integrating Information from Printed and Digital Documents

In a third study we evaluated how CoScribe supports users in integrating information that is distributed between several printed and digital documents. This setting was inspired from the findings of our third field study (see Section 2.3.3), which show that learners often work with interconnected collections of printed and digital documents.

We assessed the ease-of-use and the learnability of the interactions for creating and following cross-media hyperlinks as well as the use of the tabletop display. In addition, we had the following formal hypotheses for the experiment:

H1: The system enables users to perform a complex information retrieval task in a set of interlinked printed and web documents more quickly than with traditional pen and paper.

H2: Printed and digital documents are perceived as being more closely connected when using cross-media hyperlinks.

### 7.4.1 Method

A total of 10 psychology and 6 computer science students (9 female, 7 male) participated to 1.5 hour single-user sessions. Their average age was 25.3 years ( $SD=5.5$ ). The participants judged their computer skills as high ( $M=5.5$  on a 7-point Likert scale,  $SD=1.5$ ). They are experienced knowledge workers indicating to work 24.0 hours per week with digital documents ( $SD=10.9$ ) and 13.5 hours with paper documents ( $SD=7.6$ ). A significant amount of time ( $M=11.2$  hours,  $SD=7.9$ ) is spent on working with both types of documents in the same time. Participation to our study was voluntary and no compensation was given.

They used a digital pen, paper documents and digital documents on the pen-enabled display. Eight participants used the display in a tabletop configuration while the remaining eight participants used it as a vertical screen.

The sessions were structured as follows: After three minutes of training, the first task of the test person was to create arbitrary hyperlinks between a printed document, which was provided by us, and Web pages. The next task was to follow these hyperlinks.

The remaining time of the experiment consisted of answering questions on historic and present murder cases. For this purpose, the participant was given a collection of printed and web documents that contained information about this case. The information was collected from newspaper articles and diverse Web pages. This collection was pre-structured (annotated and interlinked) with respect to specific questions. During the experiment, we observed the participants navigating within the document collections and measured the time needed until the questions were correctly answered.

A within-subject design was used for this experiment. The participant was trained on a first document set. Two other document sets were then used for testing under either condition. We counterbalanced the document sets and the order of the two conditions. In one condition, the participant could use CoScribe. Links on printed documents could be followed by tapping with the pen on the hyperlink. Digital documents could be accessed on the pen-enabled display. In the control condition, the participants used traditional pen and paper and a traditional Web browser. The printed documents contained the same annotations and hyperlinks as in the other condition, but hyperlinks were expressed as handwritten references pointing to the number and the page of the target document. The participants were given an overview document that listed all documents, their numbers and the URLs of Web pages. References on Web pages were visualized by the Web annotation tool Diigo [Dii] and indicated the number and the page of the target document. Moreover, target passages of references were highlighted in yellow, as done automatically by the CoScribe Web browser plug-in in the other condition.

Each document collection consisted of 3 to 4 printed documents with a total of 15 to 16 A4 pages mainly consisting of text. Moreover, it included 3 up to 5 web documents from Wikipedia and online newspaper archives. The documents comprised a short biography of the murderer, detailed documents on the murder and some more documents with background information on the case. The questions asked for factual information. Nevertheless, a correct answer required to integrate information from various passages of different documents. The relevant information for answering a question was distributed between 3 to 8 passages. Each collection contained between 8 and 10 hyperlinks that connected related passages. In order to guarantee comparability across the participants and for reasons of experiment duration, we opted not to let the participants themselves prepare the document collections. Instead, we asked a student who was not familiar with our system to prepare the document collections for all participants.

The goal of this task was to find out if the participants would be able to complete a realistic information integration task more quickly using CoScribe. This task goes far beyond simply following a hyperlink (which obviously takes less time than manually searching the referenced passage). Users had to handle many documents containing not only relevant, but much irrelevant information. Moreover, the users had to decide on the relevance of hyperlinks, since only about one third of all hyperlinks linked to passages which were relevant for a specific question. Finally, the participants were pressed for time (15 min. per document set for three questions). This setting thus represented a realistic knowledge work task where a co-worker had pre-structured a collection of documents with regard to specific aspects.

After the experimental session, we gathered subjective feedback with a questionnaire. This contained 24 items covering the usability of cross-media hyperlinks and of the pen-enabled display, the perceived proximity between paper and digital documents and personal information. For the sake of readability, we will present all statements in their positive form, but the form varied in the questionnaire.

Finally, we conducted a semi-structured interview with each participant to gather

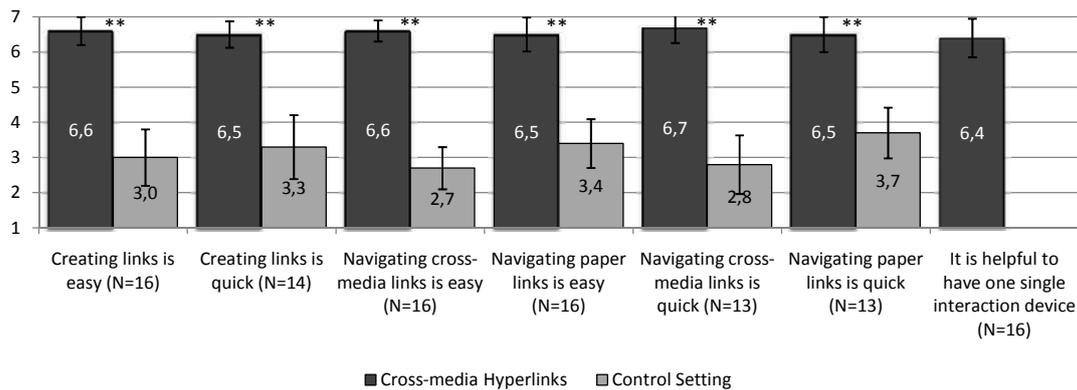


Figure 7.9: Quantitative responses for hyperlinks. (All statements are transformed to their positive form. Error lines indicate the 95 % confidence intervals of the means. \*\* = statistically significant)

qualitative feedback.

## 7.4.2 Results and Discussion

**Creating and following links.** After a few minutes of training, all users had learnt how to use the pen on the display and how to create and follow links. Responses to an open question showed that the interaction was appreciated as being “easy” (8 persons), “quick” (2 persons) and “highly intuitive” (2 persons). We observed all but one participant navigating through the interlinked documents with high confidence as if they had been using the system already for a long time (Requirements R4, R4a).

In the responses to the questionnaire, both creating hyperlinks was judged to be significantly easier and faster with the system than in the control setting (Fig. 7.9). Creating a hyperlink was rated to be very easy ( $M = 6.6$  on a 7-point Likert scale,  $SD = .8$ ,  $N = 14$ ) in contrast to the value for creating traditional references which was below average ( $M = 3.0$ ,  $SD = 1.5$ ,  $N = 14$ ) ( $T = 7.66$ ,  $df = 13$ ,  $p < .001$ ). Alike, creating a cross-media hyperlink was rated to be significantly faster ( $M = 6.5$  on a 7-point Likert scale,  $SD = .7$ ,  $N = 14$ ) than creating a traditional reference ( $M = 3.3$ ,  $SD = 1.7$ ,  $N = 14$ ) ( $T = 7.19$ ,  $df = 13$ ,  $p < 0.01$ ).

In the interviews, the users valued the tap-and-hold gesture for starting hyperlinks, as this permits a clear differentiation between annotations and the link command. As shown above, the gesture was considered to be very easy and quick to perform. This stands in contrast to the findings of Li et al. [LHGL05] which report higher error rates and larger task completion times with hold-down gestures. This contradiction might relate to the fact that Li et al. used tap-and-hold for mode switching in a drawing application, where mode switching occurs very frequently.

In contrast, creating links is an activity that is performed only at an infrequent basis. In this latter case, holding down the pen for half a second is much more acceptable. As indicated by several participants, this short interruption is even supportive because it underscores that a special mode for creating (instead of following) a hyperlink is now being entered.

The separate hyperlink column was appreciated for enabling a quick overview of all links. However, 4 participants desired the possibility to also create hyperlinks on individual words directly within the document.

As far as following links is concerned, this was also rated to be significantly quicker and faster than in the control condition (R4a) (see Fig. 7.9). The difference is particularly pronounced for cross-media hyperlinks between a printed document and a Web page, but also for hyperlinks between two paper page, the difference is considerable. The detailed values are as follows: Following cross-media hyperlinks was judged to be very easy with CoScribe ( $M = 6.6$  on a 7-point Likert scale,  $SD = .6$ ,  $N = 16$ ), whereas the value is below average in the control setting ( $M = 2.7$ ,  $SD = 1.2$ ,  $N = 16$ ) ( $T = 10.7$ ,  $df = 15$ ,  $p < .001$ ). Following hyperlinks between two paper documents was judged to be slightly less easy ( $M = 6.2$ ,  $SD = 1.2$ ,  $N = 14$ ), but still much easier than without CoScribe ( $M = 3.4$ ,  $SD = 1.9$ ,  $N = 14$ ). Again, this difference is statistically highly significant ( $T = 4.26$ ,  $df = 13$ ,  $p = .001$ ). A cross-media hyperlink can be followed much more quickly ( $M = 6.7$ ,  $SD = .8$ ,  $N = 13$ ) than a traditional reference ( $M = 2.8$ ,  $SD = 1.5$ ,  $N = 13$ ) ( $T = 5.31$ ,  $df = 12$ ,  $p < .001$ ). This difference is slightly less pronounced, but still significant, for links between two paper documents with CoScribe ( $M = 6.5$ ,  $SD = .9$ ,  $N = 13$ ) and in the control setting ( $M = 3.7$ ,  $SD = 1.3$ ,  $N = 13$ ) ( $T = 5.31$ ,  $df = 12$ ,  $p < .001$ ).

It was considered very helpful to have the same interaction device for printed and digital documents ( $M = 6.4$  on a 7-point Likert scale,  $SD = 1.1$ ,  $N = 16$ ). Participants reported that this makes the interaction faster and connects printed and digital documents more directly (R4). In the interviews, 14 participants valued that one does not have to switch between different devices for paper and for the computer screen. One participant emphasized on the fact that using the same pen for creating a hyperlink between paper and digital documents is helpful, as it gives the feeling of connecting documents more directly.

**Tabletop setting.** When comparing the pen-enabled tabletop display with the pen-enabled vertical screen, the tabletop configuration was clearly preferred. While four participants reported in the interviews to prefer a vertical screen because it is easier to read and printed documents cannot cover parts of the display, ten participants preferred a horizontal tabletop configuration. Most important reasons mentioned in the interviews were first that it is more natural and ergonomic to use a pen on a horizontal surface and second that printed and digital documents are more closely coupled using one surface for both of them (R1a). However, three participants perceived an extra effort for rearranging documents on the display. This discomfort is due to two deficiencies of our current prototype: the space provided on the tabletop

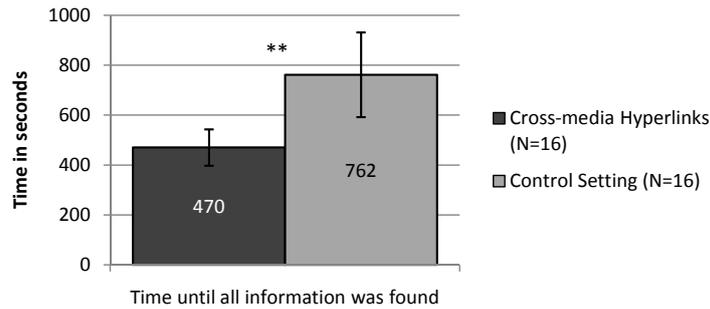


Figure 7.10: Task completion times. (Error lines indicate the 95 % confidence intervals of the means. \*\* = statistically significant)

is limited and no provision is made for coping with occlusions of displayed pages.

A further problem of the current tabletop prototype is the image quality of the display. While this is comparable to the quality of other tabletop displays, it does not yet reach the quality of standard screens. This is not problematic in many application areas of tabletop displays, for example interactive exhibits in museums. However, knowledge work has higher requirements concerning display quality. Both contrast and luminance of our current prototype are acceptable for using the display for some hours but must be improved if the display shall be used by knowledge workers for entire workdays.

When asked about the convenience of using the pen on the screen, the participants that had used the screen in a tabletop configuration reported a better score ( $M = 6.3$ ,  $SD = .7$ ,  $N = 8$ ) than those of the vertical display condition ( $M = 5.6$ ,  $SD = 1.4$ ,  $N = 8$ ). While this difference is not significant, it is in line with our observations that it was more difficult for the participants to use the pen on the vertical display.

**Performance gain (H1).** As depicted in Fig. 7.10, completing the information answering task with CoScribe took in average only about 60 % of the time needed in the control setting. This difference is highly significant ( $T = -3.22$ ,  $df = 15$ ,  $p < .01$ ) and confirms our hypothesis H1. As the two document collections did not result in significantly different completion times, we analyzed them together.

We observed that it is very important to label a link anchor with some information about the target document (such as “Biography of the murderer”). While most links were labeled, each document collection contained the same number of unlabeled links. With the latter ones, the participants were much more likely to get disoriented and to be uncertain if they had already followed these links.

**Gap between printed and digital documents (H2).** In the questionnaire, we asked the participants as how closely connected they perceive printed and digital docu-

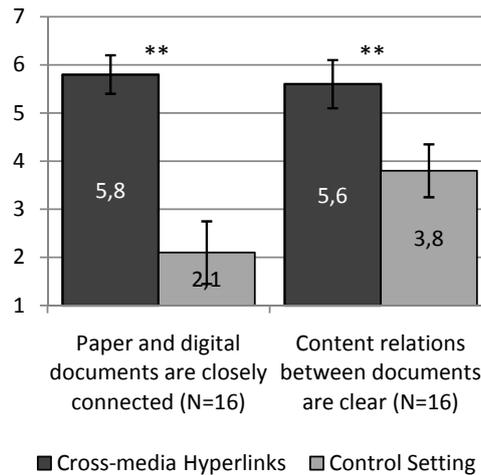


Figure 7.11: Closeness of printed and digital documents. (Error lines indicate the 95 % confidence intervals of the means. \*\* = statistically significant)

ments. Figure 7.11 shows that with cross-media hyperlinks and the pen-enabled display, printed and digital documents are considered to be significantly more closely connected ( $M = 5.8$  on a 7-point Likert scale,  $SD = .8$ ,  $N = 16$ ) compared to the control setting ( $M = 2.1$ ,  $SD = 1.3$ ,  $N = 16$ ) ( $T = 9.8$ ,  $df = 15$ ,  $p < .001$ ). Alike, the relations between the contents of printed and digital documents are found to be closer with CoScribe ( $M = 5.6$ ,  $SD = 1.0$ ,  $N = 16$ ) than in the control setting ( $M = 3.8$ ,  $SD = 1.1$ ,  $N = 16$ ) ( $T = 5.2$ ,  $df = 15$ ,  $p < .001$ ).

We finally asked the participants on the likelihood that they would use CoScribe if it was available to end-users. The average response elevated at  $M = 5.8$  ( $SD = 1.2$ ,  $N = 16$ ). This is supporting evidence for a high acceptability of CoScribe.

## 7.5 Conclusions

The results of three user studies of CoScribe indicate that the integration of paper and digital documents in seamless interaction techniques and in a coherent and consistent system enhances both work performance and user satisfaction. Overall, the results support the assumption of a *high acceptability* of the system. This indicates that CoScribe is likely to be accepted and used in realistic settings. Moreover, the studies showed that the new interactions are simple, easy to learn and reliable. For a question answering task with interconnected document collections, we could further demonstrate a *significant performance gain* when using CoScribe. As shown in one of our introductory field studies, this setting, in which users integrate information from hybrid collections of printed and digital documents, is frequent in knowledge work.

CoScribe proved to be far less distractive for making annotations and classifying them than a laptop but *compares favorably with the ease-of-use of traditional pen and paper*. Tag Buttons for classifying annotations were used to a considerable degree. In terms of use, the category “important” largely outperformed all other categories with more than 10 % of all annotations being classified with that category. Users valued that Tag Buttons are quick and easy to use and do not interrupt the main task of annotating. However, the lack of real-time feedback on the current classification mode and on the automatic clustering of annotations turned out to be problematic, in particular for privacy-critical tasks such as defining the visibility of an annotation. Currently, there exist no Anoto pens that can provide system feedback while supporting real-time streaming of pen data. It can be assumed that this will change in the near future. A next version of CoScribe should use these pens and provide real-time feedback on the current classification mode, for example by indicating it on a small display directly on the pen. Moreover, in addition to the classification of annotations, it should be possible for the user to define a precise area on the document where a classification applies to. This fall-back to a purely syntactic way of classification, which does not involve any possibly error-prone computer interpretation, would further enhance the reliability of the interactions in settings where no system feedback is available. Further, many participants expressed the wish of switching between different pen tips for writing in different colors with one single Anoto pen. This would require developing new hardware.

The *multi-user view*, which integrates one’s own and shared annotations, was judged to be very helpful both for getting an overview on all shared annotations and for finding specific annotations. The data gathered during the experiment indicates that with the multi-user view, it takes less time to find shared annotations than with the single-user views. *Digital Paper Bookmarks* proved to be very easy to create and were judged helpful for finding pages in printed and in digital documents. A more thorough comparison with other types of tagging should be made in a follow-up study.

The results of the third study on *cross-media hyperlinks* showed that the interactions for creating and following hyperlinks are easy and reliable. A complex information integration task with hybrid collections of printed and digital documents could be performed in significantly less time using CoScribe than in the control condition. This task represents many settings in knowledge work in which information must be integrated from various sources.

Our main goal of cross-media hyperlinks and the pen-based display was to *bridge the gap* between printed documents (e.g. PDF documents, books) and digital documents (e.g. Web pages). The results of the study indicate that this goal was reached. The participants perceived printed and digital documents as being significantly closer connected than in the control setting. Moreover, also the relations between the contents of printed and digital documents were judged to become clearer with CoScribe. The participants valued that one can interact with one single pen on both types of media. However, the image quality of the current prototype of the pen-enabled display needs to be improved to be usable during an entire workday.

All studies demonstrated that the interaction techniques can be *efficiently utilized even by novices* after a few minutes of training. In the questionnaire and interviews, the interactions were judged to be easy to learn and easy to use. Our observations of the participants confirmed these subjective ratings. Three to five minutes of training were sufficient in all studies to enable the participants to use the interactions without difficulties. This is supporting evidence for the theoretical interaction model which is the foundation of CoScribe's interaction techniques. Even though CoScribe offers comprehensive functionality, the Pen-and-Paper User Interface remains easy to learn and easy to use because it is based on a small amount of simple core interactions and draws its inspiration from well-established practices of interacting with pen and paper.

On average, the participants indicated that it is very likely that they would use CoScribe in its current state if it was available for regular use. This underscores again the high degree of acceptability encountered in the three studies. Based on these findings, follow-up studies should be conducted in which knowledge workers use CoScribe during their normal work for longer periods of time, overcoming the limitations of short-term user studies. This could provide deeper insights on how CoScribe affects document use and collaborative practices.

# 8

## Conclusions

Highly varied and efficient practices of interacting with physical documents have evolved over hundreds of years. Compared to these practices, the way we work today with documents on computers is, in many respects, a step backwards. It is therefore an important advance to extend digital interaction from the desktop into the physical space. However, this presents enormous challenges for designing the user interface.

The goal of this thesis was to provide more advanced support for knowledge workers who use both paper and digital documents. Based on an observation-driven and empirically-inductive approach, we aimed at developing new interaction techniques which offer richer interactions than previous work and which combine paper and screens more closely while maintaining the full mobility of paper. These techniques had the goal to provide for both co-located and remote collaboration and to offer comprehensive support instead of supporting only a small fraction of the practices which are frequent in information ecologies.

In this final chapter, we summarize the main outcomes of this thesis (Section 8.1) and identify directions of future research (Section 8.2).

### 8.1 Summary

In this thesis, we take the interaction with printed and with digital documents a significant step further. This is particularly related to the following points.

**Initial field studies and ecological perspective.** Three field studies provided a solid grounding for the theories and concepts developed in this thesis. They allow a thorough insight into how students use pen and paper and digital media in their daily work at universities. We found that *pen and paper is highly important* in that domain with a large proportion of students taking their notes only with traditional

pen and paper and not with digital technology. However, we also observed that many students *tightly combine pen and paper with digital media*. This was particularly pronounced with some learning group meetings in which the participants switched from printed documents to documents on a laptop and vice versa during one single task. The study further showed that collaboration and interaction with other people is inherently linked to interaction with documents, such as writing, pointing or handing over a document. For this reason, a document-centered approach seems to be appropriate for encompassing collaboration support. The tight coupling of interaction with physical and digital artifacts as well as with collaborators, which we have observed in practice, brought us to applying an *ecological perspective* as the guiding theoretical basis of this thesis. This follows the concepts of distributed cognition and information ecologies. Prior research on systems that integrate paper and digital information lack a theoretical basis of interaction. As we will discuss below, the *explicit theoretical grounding* of our work gave rise to a broader support and to more advanced solutions than in related work.

**Interaction model.** Based on the field studies and on the ecological perspective, we developed a *theoretical model of interaction with Pen-and-Paper User Interfaces*. This is an important contribution, as there exists no previous work which models paper-based user interfaces from a theoretical perspective on interaction. By analytically separating a semantic from a syntactic level of interaction, we identified generic interaction primitives of paper-based user interfaces. The model is both of *analytical value* and provides *guidance for the design* of rich and complex interfaces which are nevertheless reliable and easy to use. Instead of mimicking interactions of Graphical User Interfaces, the interaction is geared to the *varied traditional practices* of using pen and paper and leverages the *rich interactions* that are made possible by combining multiple paper surfaces. While we show that the model can be used to describe systems from previous work, the best proof of the applicability are the novel interaction techniques developed in this thesis. Based on the model, we presented *highly versatile and intuitively usable interaction concepts* that offer a rich user experience being based on a set of tools which are made out of paper.

**CoScribe.** We contributed CoScribe, a system framework for collaborative paper-based knowledge work. Whereas previous solutions supported only a small fraction of the practices which are frequent in typical information ecologies, CoScribe provides more *comprehensive functionality*: With annotating, hyperlinking and tagging, it supports three central and generic activities of working with documents that complement each other. Moreover, it enables *various forms of collaboration* around the rather static medium of paper supporting both co-located collaboration with multiple users and remote collaboration in settings ranging from a small number of users up to several hundreds of users. Thereby CoScribe closely combines multiple documents, users and processes in one integrated environment. This is a result of the theoretical grounding on the ecological viewpoint. An ecological visualization concretizes this viewpoint for the end user and provides an integrated view of

documents, links, tags and users. To the knowledge of the author, CoScribe is the currently most comprehensive concept and realization of a technical system that integrates printed and digital documents.

As shown by the evaluation, CoScribe significantly narrows the gap between printed and digital documents by offering the *same device for interacting with paper and with screens*. Previous work either requires different interaction devices for paper and screens or restricts the use of paper to a small, digitally augmented surface. In contrast, all functions of CoScribe can be controlled with the same digital pen and the full mobility of paper is maintained.

**Flexible layout of the printed UI.** With respect to the annotation of printed documents, CoScribe covers an aspect which is highly important for efficient use and which has not been addressed by previous work: the layout of the printed documents. Our evaluation studies show that depending on personal preferences and the context, users prefer *individual print layouts*. This is supported by CoScribe. As the model abstracts annotations from the layout of the document, annotations can be shared with other persons who use different layouts and can be re-printed in a different layout with all annotations correctly resized and repositioned. Moreover, the user can *dynamically create arrangements of pages* depending on the current needs. These two means provide for a printed user interface that has a much higher degree of flexibility than offered by previous work for annotating paper documents.

**Paper-based sharing.** To the best of our knowledge, CoScribe is the first concept that comprehensively addresses the challenges of the asynchronous sharing of annotations which are made on printed documents. This comprises both sharing annotations on paper and visualizing shared annotations as well as the important aspects of privacy and access control. We have presented a *paper-based mechanism for classifying and sharing annotations*, which is seamlessly integrated with annotating and notetaking. This provides for flexibly controlling other persons' access to one's own annotations on various levels of visibility. The evaluation shows that this is easy to learn and can be efficiently used in a regular lecture after only some minutes of training. A current limitation is the small amount of feedback available in mobile settings where only a digital pen and paper is used. In the near future, novel generations of digital pens which include a display can overcome this limitation.

As for the *access to shared annotations*, the small number of previous work addressing this issue by separating the annotations of multiple users into different views, creating individual document pages for each user. This conceptually simple approach is somewhat counterintuitive, as it forces users to access several different instances of the same page for accessing all annotations that other users have made on one single document page. In contrast, our approach keeps closer to the printed document, as one printed page is represented by one single digital instance. This *novel collaborative visualization* integrates both one's own annotations and shared ones of other users. The problem of the restricted page size is addressed by a dynamic scaling approach that enlarges and shrinks individual annotations. As shown

in a user study, this is particularly *supportive for gaining an overview* on all annotations and *for seeking specific annotations*.

**Cross-media hyperlinks.** A direct consequence of the ecological viewpoint is that CoScribe offers support for user-creatable hyperlinks. These enable to integrate multiple documents instead of working only with single documents. We contribute a novel *mechanism for creating and following hyperlinks* between any combination of printed documents, their digital representations and Web pages. This puts strong emphasis on the associative character of hyperlinks, as a bidirectional hyperlink is created by an associative pen gesture that connects both link anchors. Links can be followed by tapping with the same pen on a link hot-spot either on paper, in the CoScribe viewer or in the Web browser. While similar gestures have been discussed by previous work, our concept uses *one single pen for all activities*, maintains the *full mobility of paper* and provides for *very flexible link anchors*. These latter range from individual lines of text over entire documents up to collections of documents. This is the first solution for creating own hyperlinks between printed documents and Web pages, which is an important direction for extending the hyperlink principle from the Web to physical paper. A user study showed that hyperlinks between printed documents and Web pages lead to a *significantly higher performance* in information integration tasks with hybrid document collections than traditional references.

**Tagging of documents and processes.** In addition to mechanisms for annotating and linking documents, we presented an extensive set of novel tagging techniques. These cover not only the dimension of *tagging documents*, which is well-known in knowledge work and personal information management, but we open up the novel class of *tagging temporal processes* with semantic information. The richness of the interaction techniques, which are based on our theoretical interaction model, becomes particularly noticeable here. A first technique leverages *tangible stickers* for tagging pages of physical documents. As we have seen, this is highly intuitive, provides full visual feedback on paper and affords quick access to a document page not only using a computer but also using the printed document instance. However, the tags are rather course-grained as they always apply to entire pages. A second technique, which offers very flexible scopes, uses separate *paper cards* for defining, applying and sharing freely-chosen tags. By introducing with the paper card the concept of a separate paper “tool”, we provide physical instantiations of the tag sets, which remain inaccessible in previous paper-based systems. These instantiations stimulate the convergence of tags, operations on the entire tag set and sharing of tag sets. A third technique aims at being particularly fast to be easily integrated into notetaking tasks. This is based on predefined categories that can be applied by tapping on specific *buttons that are printed on paper documents*. Finally, we discussed the novel class of *process tagging with tangible tools*. We contribute two first prototypes, the process cube and the process knob, that illustrate this principle. The *variety of interaction metaphors* for tagging stands in contrast to previous work on paper-based tagging which relies only on pen gestures.

## 8.2 Directions of Future Research

In Chapter 7, we already pointed out that future research should examine the use of the novel interaction techniques over longer periods of time. This provides additional insights into how CoScribe affects document use and collaborative practices. On a more general level, we moreover identified promising directions of research for developing further the concepts created in this thesis and the insights obtained. These directions are as follows.

### Occlusion-aware Visualization

Traditionally, the surface of a screen is exclusively reserved for displaying digital contents. The visualization does not have to cope with occlusions, as no objects are placed on the screen, not even the user's hands because the interaction is indirected by a mouse and a keyboard. In contrast, our approach closely couples printed and digital documents by combining them on one single tabletop surface. Placing *physical paper* sheets on top of the display leads to *occlusions of the digital contents*. In our current prototype, users must manually rearrange physical documents and windows on the screen to avoid occlusions.

One area of future research concerns visualization and interaction strategies for hybrid physical and digital display surfaces on tabletops. Very recently, first work has been published that addresses occlusions caused by the user's hands during pen and touch interaction [BLS<sup>+</sup>09]. The positions of the hands are tracked and menus are relocated accordingly to minimize occlusions. However, this setting differs from hybrid physical and digital display surfaces for several reasons. First, paper documents may occlude much larger areas of the display than hands. Second, the occlusions are not temporary but may persist over long periods of time. Finally and most important, it is not clear *per se* if digital documents which are occluded by paper should be relocated or, in contrast, are precisely intended to be hidden. This depends on the value that a document currently has for the user. If the digital document is in the focus of the user and becomes occluded by a physical page, it should certainly be relocated to remain visible to the user. However, if the document is currently not in the user's focus and a physical page is placed on top of it, the user might consciously cover it, for example by creating a hybrid pile consisting of digital and printed documents.

Future research should address the non-trivial issue of *distinguishing between intended and unintended occlusions*. Moreover, it should examine appropriate *strategies for automatic rearrangements of digital contents* as well as *techniques for interacting with hybrid piles* of printed and digital documents.

### ePaper Displays

Our approach brings information on paper very close to complementing digital information because paper documents are used on the same surface as digital docu-

ments and the interaction relies on the same digital pen. However, if digital information shall be updated directly on a paper document, this needs to be reprinted. An area of future research consists of enabling *continuous updates of information on paper* in real time.

A concept which has already been explored in the past is to overlay paper with digital information, for example by projecting digital contents onto paper or by placing a small display on the paper document. However, these approaches either restrict the mobility of paper (e.g. [Wel93]) or overlay only a small area of the paper document with digital information (e.g. [MPL<sup>+</sup>02, SGF<sup>+</sup>09]). Both approaches moreover restrict the natural interaction with paper, as paper sheets have to be kept flat.

A highly promising development is ePaper [CP08]. This is a novel technology for displays that are very thin, flexible and deformable, provide for a very high contrast, are well readable even in difficult light conditions and consume little energy. There already exist commercial models that allow pen interaction. However, the update rate is currently too low for interactive applications. In the future, ePaper has the potential to replace traditional paper to a much higher extent than current digital technologies. Our vision consists of an environment, in which the user disposes of a *large number of thin and malleable ePaper displays*, which behave very much like paper. In contrast to current displays, multiple ePaper surfaces provide for the interactions of laying out pages in the physical space, comparing pages and arranging them in flexible ways. As pointed out in [HVA<sup>+</sup>05], pages of digital documents would be temporarily bound to ePaper displays.

In the context of this thesis, it is an important question *to what extent and how our interaction techniques can be transferred to such an ePaper environment*. This should be examined in future work. We assume that the sharing and classification of annotations as well as the multi-user visualization of shared annotations can be transferred in a straightforward manner. In particular, all interactions that rely on associations between multiple page surfaces (e.g. linking, Menu Card tagging) make sense in a multi-page ePaper environment as well. Hence, a large proportion of our techniques generalizes not only to paper but to paper-like digital surfaces.

A significant difference between paper and ePaper is that information on paper is fix while it is only temporarily bound to an ePaper display. This provides for more flexibility of use. However, if the user disposes only of a restricted number of ePaper displays, she is forced to successively use one display for different documents or pages. This challenges tangible bookmarking, as physical bookmarks would have to automatically jog out of the paper display. For all kinds of physical indexing, it would be highly supportive to develop *ePaper displays that are able to change their physical shape*. A second challenge are filing interactions, such as placing a document into a folder. Traditional paper would remain in the folder and thereby physically indicate its contents, the number of documents filed etc. If in contrast, the ePaper display is reused and does not remain within the folder, this physical information gets lost. Future research should therefore not only address the question of *how a set of ePaper displays can be quickly bound to a set of documents* but also the issue of *how*

*interaction techniques and tangible tools can compensate for the physical information that gets lost when removing the unambiguous mapping between a document page and its physical carrier surface.*

### **Support of Further Core Interactions**

We deliberately decided not to augment paper by electronic components. This allowed to maintain paper a cheap medium which digital contents can be easily printed onto. A drawback of this approach is that some interactions with paper cannot be digitally tracked. Our core interactions rely on writing with the pen on paper and on changing the physical location of paper pages. For example *bending or folding pages* is not comprised.

The use of ePaper displays – as discussed above – would make our premise obsolete, since the user would not print contents on paper but would use a technological product. This can be enhanced with sensors for additional core interactions without drawbacks for everyday use. This opens up an area of future research that examines *further core interactions to be performed with paper-like surfaces.*

### **Task Support**

Two final directions of future research concern supporting the tasks that are performed in a paper-based environment.

This thesis focused on tasks that are centered around the reading of documents, such as annotating, integrating and abstracting. While this is accompanied to some extent by writing, our focus is not the task of creating new documents. Paper-based *support for collaboratively writing documents* is a promising direction of research. While it is clear that the actual redaction of a document can be better performed using word processing than pen and paper, paper can provide good support for the stages that precede and accompany the actual redaction. This includes creativity tools for the beginning of the process, information integration tools for literature search, structuring tools for defining the outline of the document and tools for proof-reading (e.g. [WISN08]). With support for reading and integration, CoScribe assists in partial aspects of this process.

Second, the paper-based infrastructure can be used to support knowledge work by *automatic analyses*. As far as only digital contents are concerned, this is an active area of research. Aggregations of user generated content in the Web 2.0, recommender systems and intelligent tutoring systems are only a few examples thereof. By extending these principles to physical documents and to physical interactions, computer systems will support interaction with documents to a more comprehensive extent. In the context of our interaction techniques, this could for example comprise automatically aggregating shared annotations, links and tags and recommending relevant annotations, documents and passages. This will take us another step further on the challenging and exciting way towards closing the gap between paper and digital information.



# List of Figures

|      |  |    |
|------|--|----|
| 1.1  | Contributions of this thesis . . . . .                                       | 5  |
| 1.2  | Schematic overview of the research method used in this thesis . . . . .      | 7  |
| 1.3  | Temporal overview on the phases of the research process . . . . .            | 8  |
| 1.4  | Structure of this thesis . . . . .   | 9  |
|      |  |    |
| 2.1  | Two-dimensional space of learning forms . . . . .                            | 15 |
| 2.2  | Context factors in notetaking assessed in the second field study . . . . .   | 21 |
| 2.3  | Combinations of media used to take notes on . . . . .                        | 23 |
| 2.4  | Use of course material . . . . .   | 25 |
| 2.5  | Advantages of notetaking media . . . . .                                     | 26 |
| 2.6  | Follow-up activities of notetaking . . . . .                                 | 27 |
| 2.7  | Screenshot of the tool for textual annotations and highlightings . . . . .   | 29 |
| 2.8  | Average number of annotations made by each participant per session . . . . . | 31 |
| 2.9  | Average number of annotations per user of each session . . . . .             | 32 |
| 2.10 | Examples of handwritten annotations . . . . .                                | 33 |
| 2.11 | Types of handwritten annotations . . . . .                                   | 34 |
| 2.12 | Media use in a typical learning group meeting . . . . .                      | 36 |
| 2.13 | Main elements of an information ecology . . . . .                            | 41 |
|      |  |    |
| 3.1  | Linear and two-dimensional barcodes . . . . .                                | 56 |
| 3.2  | Desktop scanner . . . . .  | 57 |
| 3.3  | Scanning pens . . . . .  | 58 |
| 3.4  | Schematic view of the Digital Desk . . . . .                                 | 59 |
| 3.5  | Pen-based notetaking tools . . . . .   | 60 |
| 3.6  | Anoto pattern . . . . .  | 61 |
| 3.7  | Anoto pens . . . . .   | 62 |
| 3.8  | IDEas notebook . . . . .   | 65 |
| 3.9  | Interaction Lens of the A-Book . . . . .                                     | 66 |
| 3.10 | ModelCraft . . . . .   | 68 |
| 3.11 | Designer's Outpost and PaperPoint . . . . .                                  | 69 |
| 3.12 | WebStickers . . . . .  | 70 |

|      |   |     |
|------|---|-----|
| 3.13 | Pre-programmed hyperlinks . . . . .   | 71  |
| 3.14 | PaperLink . . . . .   | 73  |
| 3.15 | PapierCraft . . . . .   | 74  |
| 3.16 | PaperPDA . . . . .  | 75  |
| 3.17 | DocuDesk . . . . .  | 76  |
| 3.18 | Tagging gestures of PapierCraft . . . . .   | 78  |
| 3.19 | Overview of the applications in the design space: conceptual activities and collaboration . . . . . | 81  |
| 3.20 | Perceptual coupling of paper and digital documents . . . . .  | 82  |
| 4.1  | Basic PPUI . . . . .  | 89  |
| 4.2  | Advanced PPUI of CoScribe . . . . .   | 90  |
| 4.3  | The analytical framework applied to CoScribe . . . . .  | 91  |
| 4.4  | Distinctive features of the two principles . . . . .  | 101 |
| 4.5  | The two principles of how information is related between different representations . . . . .        | 102 |
| 4.6  | Time/space matrix of collaboration . . . . .  | 105 |
| 4.7  | Collaborative settings . . . . .  | 106 |
| 5.1  | Overview of CoScribe . . . . .  | 112 |
| 5.2  | Anoto pen . . . . .   | 114 |
| 5.3  | CoScribe viewer . . . . .   | 115 |
| 5.4  | Co-located collaboration . . . . .  | 117 |
| 5.5  | Remote collaboration . . . . .  | 117 |
| 5.6  | Using the same digital pen to interact on real paper and on displays . . . . .                      | 120 |
| 5.7  | Tangible tools which are made of paper . . . . .  | 122 |
| 5.8  | CoScribe Firefox plug-in . . . . .  | 123 |
| 5.9  | The ecological view . . . . .   | 125 |
| 5.10 | Filter options of the ecological view . . . . .   | 127 |
| 5.11 | Printed buttons . . . . .   | 128 |
| 5.12 | Example of an annotation which includes sketches . . . . .  | 130 |
| 5.13 | Annotating a printed document . . . . .   | 132 |
| 5.14 | A document containing printed annotations . . . . .   | 133 |
| 5.15 | Example layouts of printouts . . . . .  | 136 |
| 5.16 | Printed button toolbar . . . . .  | 138 |
| 5.17 | Pen feedback . . . . .  | 139 |
| 5.18 | The CoScribe document viewer . . . . .  | 141 |
| 5.19 | A printout including annotations of two users . . . . .   | 142 |
| 5.20 | Example of an annotation with text which is difficult to recognize . . . . .                        | 146 |
| 5.21 | An annotation that combines text with graphics . . . . .  | 147 |
| 5.22 | Pen gestures for creating a hyperlink . . . . .   | 154 |
| 5.23 | Details of the pen gestures . . . . .   | 154 |
| 5.24 | Types of association areas . . . . .  | 156 |
| 5.25 | Link gesture on physical books . . . . .  | 157 |

---

|      |  |     |
|------|--|-----|
| 5.26 | Physical folder . . . . .  | 158 |
| 5.27 | Activating a hyperlink . . . . .   | 159 |
| 5.28 | Digital Paper Bookmarks . . . . .  | 164 |
| 5.29 | Bookmarking technique . . . . .  | 165 |
| 5.30 | Semantic types of bookmarks . . . . .  | 166 |
| 5.31 | Visualization of bookmarks in the CoScribe viewer . . . . .                          | 168 |
| 5.32 | Collaborative bookmark view . . . . .  | 169 |
| 5.33 | Tag Menu Card . . . . .  | 170 |
| 5.34 | Interaction technique for Tag Menu Cards . . . . .                                   | 171 |
| 5.35 | Process Cube . . . . .   | 174 |
| 5.36 | Exemplary model of semantic phases for structuring learning group meetings . . . . . | 175 |
| 5.37 | Process Knob . . . . .   | 175 |
| 5.38 | Mapping of CoScribe's main concepts to the ecological perspective . . . . .          | 177 |
|      |  |     |
| 6.1  | Overview on the software and hardware components of CoScribe . . . . .               | 182 |
| 6.2  | Example of physical and logical pages . . . . .                                      | 182 |
| 6.3  | Two prototypes of pen-enabled tabletop displays . . . . .                            | 184 |
| 6.4  | Setup of the display prototype . . . . .   | 185 |
| 6.5  | Schematic illustration of the technical setup . . . . .                              | 185 |
| 6.6  | Performance of the algorithm for clustering annotations . . . . .                    | 189 |
|      |  |     |
| 7.1  | Average distraction level . . . . .  | 195 |
| 7.2  | Annotations per user . . . . .   | 196 |
| 7.3  | Annotations per slide . . . . .  | 196 |
| 7.4  | Percentage of semantic categories and visibilities . . . . .                         | 197 |
| 7.5  | Multi-user view and single-user view . . . . .                                       | 200 |
| 7.6  | Task completion times . . . . .  | 200 |
| 7.7  | Quantitative results for Digital Paper Bookmarks . . . . .                           | 201 |
| 7.8  | General questions on CoScribe . . . . .  | 202 |
| 7.9  | Quantitative responses for hyperlinks . . . . .                                      | 205 |
| 7.10 | Task completion times . . . . .  | 207 |
| 7.11 | Closeness of printed and digital documents . . . . .                                 | 208 |



# List of Tables

|     |   |     |
|-----|---|-----|
| 2.1 | Participants of the study . . . . .   | 20  |
| 2.2 | Overview of the requirements identified in this chapter . . . . .   | 49  |
| 3.1 | Overview of which requirements are supported by previous work . .   | 79  |
| 4.1 | Comparison of Core Interactions . . . . .   | 95  |
| 4.2 | Core interactions of a representative set of related systems and of the interaction strategies presented in this thesis . . . . . | 96  |
| 4.3 | Appropriate complementing representations . . . . .   | 104 |
| 5.1 | Desiderata and our approaches concerning the integration of printed and digital documents . . . . .                               | 119 |
| 5.2 | Desiderata and our approaches concerning the annotation of printed and digital documents . . . . .                                | 131 |
| 5.3 | Baseline performance of the handwriting recognition of lecture annotations . . . . .  | 146 |
| 5.4 | Performance of the handwriting recognition for domain-specific terms  | 148 |
| 5.5 | Desiderata and our approaches concerning cross-media hyperlinks .   | 151 |
| 5.6 | Desiderata and our approaches concerning cross-media hyperlinks .   | 163 |
| 5.7 | Comparison of interaction techniques for tagging . . . . .  | 176 |
| 7.1 | Overview of the evaluation of the requirements from Chapter 2 . . .   | 192 |



# Bibliography

- [AAH97] Toshifumi Arai, Dietmar Aust, and Scott E. Hudson. Paperlink: a technique for hyperlinking from real paper to electronic content. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 327–334, New York, NY, USA, 1997. ACM.
- [AD72] M. J. Adler and C. Van Doren. *Howe to Read a Book. Revised Edition*. Simon and Schuster, 1972.
- [AGH<sup>+</sup>98] Annette Adler, Anuj Gujar, Beverly L. Harrison, Kenton O'Hara, and Abigail Sellen. A diary study of work-related reading: design implications for digital reading devices. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 241–248, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.
- [AKM07] Erwin Aitenbichler, Jussi Kangasharju, and Max Mühlhäuser. MundoCore: A Light-weight Infrastructure for Pervasive Computing. *Pervasive and Mobile Computing*, 2007.
- [Anoa] Anoto Forms Processing. Anoto AB. <http://www.anoto.com/forms-processing.aspx>.
- [Anob] Anoto. Digital pen and paper technology. Anoto AB. <http://www.anoto.com>.
- [Aus62] John Langshaw Austin. *How To Do Things With Words*. Harvard University Press, 1962.
- [Bai04] William Sims Bainbridge, editor. *Berkshire Encyclopedia of Human-Computer Interaction*. 2004.
- [BCG<sup>+</sup>01] Maribeth Back, Jonathan Cohen, Rich Gold, Steve Harrison, and Scott Minneman. Listen reader: an electronically augmented paper-based book. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 23–29, New York, NY, USA, 2001. ACM.

- [Bee] BeeTagg Multicode System. Connvision Ltd.. <http://www.beetagg.com>.
- [Büh78] Karl Bühler. *Die Darstellungsfunktion der Sprache*. Ullstein, Frankfurt, 1978.
- [BHE09] Leah Buechley, Sue Hendrix, and Mike Eisenberg. Paints, paper, and programs: First steps toward the computational sketchbook. In *Proceedings of the Third International Conference on Tangible and Embedded Interaction (TEI'09)*, pages 9–12, 2009.
- [BHG<sup>+</sup>07] Florian Block, Michael Haller, Hans Gellersen, Carl Gutwin, and Mark Billinghurst. Voodosketch: Physical interface palettes and sketched controls alongside augmented work surfaces. In *Adjunct. Proceedings of UbiComp 07*, pages 789–793, 2007.
- [BHH<sup>+</sup>07] Peter Brandl, Michael Haller, Michael Hurnaus, Verena Lugmayr, Juergen Oberngruber, Claudia Oster, Christian Schafleitner, and Mark Billinghurst. An adaptable rear-projection screen using digital pens and hand gestures. In *17th International Conference on Artificial Reality and Telexistence*, pages 49–54, 2007.
- [BK06] Aaron Bauer and Kenneth Koedinger. Pasting and encoding: Note-taking in online courses. In *Proc. 6th Int. Conf. on Advanced Learning Technologies (ICALT'06)*, 2006.
- [BL00] Michel Beaudouin-Lafon. Instrumental interaction: an interaction model for designing post-wimp user interfaces. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 446–453, New York, NY, USA, 2000. ACM.
- [BLS<sup>+</sup>09] Peter Brandl, Jakob Leitner, Thomas Seifried, Michael Haller, Bernard Doray, and Paul To. Occlusion-aware menu design for digital tabletops. In *CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 3223–3228, New York, NY, USA, 2009. ACM.
- [BP05] Francoise Boch and Annie Piolat. Note taking and learning: A summary of research. *The WAC Journal*, 16:101–113, September 2005.
- [Bus45] Vannevar Bush. As we may think. *The Atlantic Monthly*, pages 101–108, July 1945.
- [CHBM07] Marcelo Coelho, Lyndl Hall, Joanna Berzowska, and Pattie Maes. Pulp-based computing: A framework for building computers out of paper. In *The 9th International Conference on Ubiquitous Computing (UbiComp 2007)*, 2007.

- [CNS<sup>+</sup>95] J. Coutaz, L. Nigay, D. Salber, A. Blandford, J. May, and R. M. Young. Four easy pieces for assessing the usability of multimodal interaction: The care properties. In *Proceedings of INTERACT'95*, 1995.
- [Coh60] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:37–46, 1960.
- [CP08] Elise Co and Nikita Pashenkov. Emerging display technologies for organic user interfaces. *Commun. ACM*, 51(6):45–47, 2008.
- [CS91] Paul Chandler and John Sweller. Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4):293–332, 1991.
- [DBEK04] Christian Decker, Michael Beigl, Adam Eames, and Uwe Kubach. Digiclip: Activating physical documents. In *ICDCSW '04: Proceedings of the 24th International Conference on Distributed Computing Systems Workshops - W7: EC (ICDCSW'04)*, pages 388–393, Washington, DC, USA, 2004. IEEE Computer Society.
- [Dii] Diigo. Social Annotation. <http://www.diigo.com>.
- [DLC<sup>+</sup>99] Richard C. Davis, James A. Landay, Victor Chen, Jonathan Huang, Rebecca B. Lee, Frances C. Li, James Lin, III Charles B. Morrey, Ben Schleimer, Morgan N. Price, and Bill N. Schilit. Notepals: lightweight note sharing by the group, for the group. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 338–345, New York, NY, USA, 1999. ACM.
- [Dow05] Stephen Downes. E-learning 2.0. *eLearn*, 2005(10):1, 2005.
- [Dru64] Peter F. Drucker. *The concept of the corporation*. New American Library of World Literature, 1964.
- [EGR91] Clarence A. Ellis, Simon J. Gibbs, and Gail Rein. Groupware: some issues and experiences. *Communications of the ACM*, 34(1):39–58, 1991.
- [EMBW08] Katherine Everitt, Meredith Ringel Morris, A.J. Bernheim Brush, and Andrew Wilson. Docudesk: An interactive surface for creating and rehydrating many-to-many linkages among paper and digital documents. In *Proceedings of IEEE Tabletops and Interactive Surfaces 2008*, pages 25–28, 2008.
- [Fia04] Mark Fiala. Artag, an improved marker system based on artoolkit. Technical Report NRC 47166, National Research Council Canada, 2004.
- [Fis04] Kenneth P. Fishkin. A taxonomy for and analysis of tangible interfaces. *Personal Ubiquitous Comput.*, 8(5):347–358, 2004.

- [Fly] Fly Fusion Pentop Computer, leapfrog enterprises inc., <http://www.flypentop.com>.
- [FvDFH90] James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes. *Computer graphics: principles and practice (2nd ed.)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990.
- [GH06] Scott Golder and Bernardo A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.
- [Gib77] James J. Gibson. The theory of affordances. In Robert Shaw and John Bransford, editors, *Perceiving, Acting, and Knowing*. Hillsdale : Erlbaum, 1977.
- [Gla67] Anselm L. Glaser, Barney G. ; Strauss. *The discovery of grounded theory : strategies for qualitative research*. Weidenfeld and Nicolson, London, 1967.
- [Gra89] D. R. Graham. Incremental development: review of nonmonolithic life-cycle development models. *Inf. Softw. Technol.*, 31(1):7–20, 1989.
- [Gre] Greasemonkey firefox extension. <http://www.greasespot.net>.
- [Gui03] François Guimbretière. Paper augmented digital documents. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 51–60, New York, NY, USA, 2003. ACM Press.
- [HB06] Eva Hornecker and Jacob Buur. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 437–446, New York, NY, USA, 2006. ACM.
- [HHK00] James Hollan, Edwin Hutchins, and David Kirsh. Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Human-Computer Interaction*, 7(2):174–196, 2000.
- [HHT99] Jeremy M. Heiner, Scott E. Hudson, and Kenichiro Tanaka. Linking and messaging from real paper in the paper pda. In *UIST '99: Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 179–186, New York, NY, USA, 1999. ACM Press.
- [HKP08] Jan Heß, Guy Küstermann, and Volkmar Pipek. p(ersonal)remote - ein individuell gestaltbares fernbedienungskonzept. In *Conference Proceedings of Mensch und Computer 2008 (MuC'08)*, 2008.
- [HPS] HP Scanjet N9210. Hewlett-Packard Development Company, L.P. <http://www.hp.com>.

- [HRG<sup>+</sup>04] Ken Hinckley, Gonzalo Ramos, Francois Guimbretiere, Patrick Baudisch, and Marc Smith. Stitching: pen gestures that span multiple displays. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 23–31, New York, NY, USA, 2004. ACM Press.
- [HVA<sup>+</sup>05] David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus Troje, and Derek Johns. Paper windows: interaction techniques for digital paper. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 591–599, New York, NY, USA, 2005. ACM Press.
- [ISOa] ISO/IEC Standard 15420:2000 (Bar code symbology specification - EAN/UPC). International Organization for Standardization.
- [ISOb] ISO/IEC Standard 16022:2000 (International symbology specification - Data Matrix). International Organization for Standardization.
- [ISO99] International Organization for Standardization: ISO standard 13407:1999. Human-centred design processes for interactive systems, 1999.
- [IU97] Hiroshi Ishii and Brygg Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA, 1997. ACM.
- [JBY93] David H. Jonassen, Katherine Beissner, and Michael Yacci. *Structural Knowledge: Techniques for Representing, Conveying, and Acquiring Structural Knowledge*. Lawrence Erlbaum Associates, Hillsdale, 1993.
- [JGH<sup>+</sup>08] Robert J.K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. Reality-based interaction: a framework for post-wimp interfaces. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 201–210, New York, NY, USA, 2008. ACM.
- [JM09] Matthew Jarvis and Masood Masoodian. Digital management and retrieval of physical documents. In *Proceedings of the Third International Conference on Tangible and Embedded Interaction (TEI'09)*, pages 47–54, 2009.
- [JYWS07] Hao Jiang, Ron B. Yeh, Terry Winograd, and Yuanchun Shi. Digipost: Writing on post-its with digital pen to support collaborative editing tasks on tabletop displays. In *ACM symposium on User interface software and technology (UIST'07) Posters*, 2007.
- [KBNR03] Boriana Koleva, Steve Benford, Kher Hui Ng, and Tom Rodden. A framework for tangible user interfaces. In *Physical Interaction Workshop*

- on Real World User Interfaces in conjunction with Mobile HCI Conference 2003*, 2003.
- [KCR<sup>+</sup>07] Kimberle Koile, Kevin Chevalier, Michel Rbeiz, Adam Rogal, David Singer, Jordan Sorensen, Amanda Smith, Kah Seng Tay, and Kenneth Wu. Supporting feedback and assessment of digital ink answers to in-class exercises. In *Nineteenth Conference on Innovative Applications of AI*, pages 22–29, 2007.
- [KDLD07] Frederic Kaplan, Son Do-Lenh, and Pierre Dillenbourg. Docklamp: a portable projector-camera system. In *2nd IEEE TABLETOP Workshop*, 2007.
- [KGWL03] Scott R. Klemmer, Jamey Graham, Gregory J. Wolff, and James A. Landay. Books with voices: paper transcripts as a physical interface to oral histories. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 89–96, New York, NY, USA, 2003. ACM Press.
- [Kie87] Kenneth A. Kiewra. Notetaking and review: The research and its implications. *Instructional Science*, 16(3):233–249, 1987.
- [Kie89] Kenneth A. Kiewra. A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review*, 1(2):147–172, 1989.
- [Kin07] Alison King. Scripting collaborative learning processes: A cognitive perspective. In Frank Fischer, Ingo Kollar, Heinz Mandl, and Jörg M. Haake, editors, *Computer-Supported Collaborative Learning*, chapter 2, pages 13–37. Springer US, 2007.
- [KK98] M. Kobayashi and H. Koike. Enhanceddesk: integrating paper documents and digital documents. *Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific*, pages 57–62, Jul 1998.
- [KK01] José Kahan and Marja-Ritta Koivunen. Annotea: an open RDF infrastructure for shared web annotations. In *Proceedings of the 10th International World Wide Web Conference*, pages 623–632, 2001.
- [KNF<sup>+</sup>01] Scott R. Klemmer, Mark W. Newman, Ryan Farrell, Mark Bilezikjian, and James A. Landay. The designers' outpost: a tangible interface for collaborative web site design. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 1–10, New York, NY, USA, 2001. ACM Press.
- [KNF09] Hideki Koike, Wataru Nishikawa, and Kentaro Fukuchi. Transparent 2-d markers on an lcd tabletop system. In *ACM Conference on Human Factors in Computing Systems (CHI'09)*, pages 163–172, 2009.

- [KP88] Glenn E. Krasner and Stephen T. Pope. A description of the model-view-controller user interface paradigm in the smalltalk-80 system. *Journal of Object Oriented Programming*, 1(3):26–49, 1988.
- [KWI<sup>+</sup>05] Matthew Kam, Jingtao Wang, Alastair Iles, Eric Tse, Jane Chiu, Daniel Glaser, Orna Tarshish, and John Canny. Livenotes: a system for cooperative and augmented note-taking in lectures. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 531–540, New York, NY, USA, 2005. ACM Press.
- [LCC07] Wei-Chu Lai, Po-Yao Chao, and Gwo-Dong Chen. The interactive multimedia textbook: Using a digital pen to support learning for computer programming. In *IEEE International Conference on Advanced Learning Technologies (ICALT'07)*, pages 742–746, 2007.
- [Lev66] Vladimir I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10:707–710, 1966.
- [LGA<sup>+</sup>07] Chunyuan Liao, François Guimbretière, Richard Anderson, Natalie Linnell, Craig Prince, and Valentin Razmov. Papercp: Exploring the integration of physical and digital affordances for active learning. In *Human-Computer Interaction INTERACT 2007*, pages 15–28, 2007.
- [LGHH08] Chunyuan Liao, François Guimbretière, Ken Hinckley, and Jim Hollan. Papiercraft: A gesture-based command system for interactive paper. *ACM Transactions on Computer-Human Interaction*, 14(4):1–27, 2008.
- [LGL06] Chunyuan Liao, François Guimbretière, and Corinna E. Loeckenhoff. Pen-top feedback for paper-based interfaces. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 201–210, New York, NY, USA, 2006. ACM Press.
- [LHGL05] Yang Li, Ken Hinckley, Zhiwei Guan, and James A. Landay. Experimental analysis of mode switching techniques in pen-based user interfaces. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 461–470, New York, NY, USA, 2005. ACM.
- [Liv] Livescribe. <http://www.livescribe.com>.
- [LMK<sup>+</sup>06] Brian Lee, Heidy Maldonado, Scott R. Klemmer, Isabelle Kim, and Paz Hilfinger-Pardo. Longitudinal studies of augmented notebook usage informing the design of sharing mechanisms. Technical Report CSTR 2006-11 09/29/06, Stanford University, 2006.
- [Log] Logitech. io2 Digital Pen. <http://www.logitech.com>.
- [LRH00] Peter Ljungstrand, Johan Redström, and Lars Erik Holmquist. Webstickers: using physical tokens to access, manage and share bookmarks

- to the web. In *DARE '00: Proceedings of DARE 2000 on Designing augmented reality environments*, pages 23–31, New York, NY, USA, 2000. ACM Press.
- [LTZ05] Tobias Lauer, Stephan Trahasch, and Bernd Zupancic. Anchored discussions of multimedia lecture recordings. In *Proceedings of the Frontiers in Education Conference (FIE 2005)*, pages T2H–12–T2H–17, Oct 2005.
- [Mac99] Wendy E. MacKay. Is paper safer? the role of paper flight strips in air traffic control. *ACM Transactions on Computer-Human Interaction*, 6(4):311–340, 1999.
- [Mar97] Catherine C. Marshall. Annotation: from paper books to the digital library. In *DL '97: Proceedings of the second ACM international conference on Digital libraries*, pages 131–140, New York, NY, USA, 1997. ACM Press.
- [Mar98] Catherine C. Marshall. Toward an ecology of hypertext annotation. In *HYPertext '98: Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space—structure in hypermedia systems*, pages 40–49, New York, NY, USA, 1998. ACM Press.
- [MHBC08] D. McGee, X. Huang, P. Barthelmess, and P. Cohen. The neteyes collaborative, augmented reality, digital paper system. In *Proc. of IEEE Symposium on 3D User Interfaces 2008*, 2008.
- [MKS07] Motoki Miura, Susumu Kunifuji, and Yasuyuki Sakamoto. Airtransnote: An instant note sharing and reproducing system to support students learning. In *IEEE International Conference on Advanced Learning Technologies (ICALT)*, pages 175–179, 2007.
- [MM08] Pranav Mistry and Pattie Maes. Intelligent sticky notes that can be searched, located and can send reminders and messages. In *International Conference on Intelligent User Interfaces (IUI'08)*, pages 425–426, 2008.
- [MMC09] Pranav Mistry, Pattie Maes, and Liyan Chang. Wuw - wear ur world - a wearable gestural interface. In *ACM Conference on Human Factors in Computing (CHI'2009) Extended Abstracts*, 2009.
- [MPL<sup>+</sup>02] Wendy E. Mackay, Guillaume Pothier, Catherine Letondal, Kaare Bøegh, and Hans Erik Sørensen. The missing link: augmenting biology laboratory notebooks. In *UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 41–50, New York, NY, USA, 2002. ACM Press.
- [ND86] Donald Norman and Stephen Draper. *User Centered System Design: New Perspectives on Human-computer Interaction*. Lawrence Erlbaum Associates, 1986.

- [Nel65] T. H. Nelson. Complex information processing: a file structure for the complex, the changing and the indeterminate. In *Proceedings of the ACM 20th national conference*, pages 84–100, New York, NY, USA, 1965. ACM.
- [NHH06] Saeko Nomura, Edwin Hutchins, and Barbara E. Holder. The uses of paper in commercial airline flight operations. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 249–258, New York, NY, USA, 2006. ACM.
- [NInPA99] Les Nelson, Satoshi Ichimura, Elin Rønby Pedersen, and Lia Adams. Palette: a paper interface for giving presentations. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 354–361, New York, NY, USA, 1999. ACM Press.
- [nKrE03] Kaj Grønbaek, Jannie F. Kristensen, Peter Ørbaek, and Mette Agger Eriksen. "physical hypermedia": organising collections of mixed physical and digital material. In *HYPertext '03: Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*, pages 10–19, New York, NY, USA, 2003. ACM.
- [NO99] Bonnie Nardi and Vicky O'Day. *Information Ecologies: Using Technology with Heart*, chapter Chapter Four: Information Ecologies. MIT Press, 1999.
- [Nok] Nokia. SU-1B Digital Pen. <http://www.nokia.com>.
- [Nor02] Donald A. Norman. *The design of everyday things*. MIT Press, London, 2002.
- [nPSN00] Elin Rønby Pedersen, Tomas Sokoler, and Les Nelson. Paperbuttons: expanding a tangible user interface. In *DIS '00: Proceedings of the 3rd conference on Designing interactive systems*, pages 216–223, New York, NY, USA, 2000. ACM.
- [NSW06a] Moira C. Norrie, Beat Signer, and Nadir Weibel. General framework for the rapid development of interactive paper applications. In *CoPADD 2006, Workshop on Collaborating over Paper and Digital Documents*, Banff, Canada, November 2006.
- [NSW06b] Moira C. Norrie, Beat Signer, and Nadir Weibel. Print-n-link: Weaving the paper web. In *Proceedings of DocEng 2006, ACM Symposium on Document Engineering*, pages 34–43, Amsterdam, The Netherlands, October 2006.
- [OAC06] Sharon Oviatt, Alex Arthur, and Julia Cohen. Quiet interfaces that help students think. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 191–200, New York, NY, USA, 2006. ACM Press.

- [Obe03] Hartmut Obendorf. Simplifying annotation support for real-world-settings: a comparative study of active reading. In *HYPertext '03: Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*, pages 120–121, New York, NY, USA, 2003. ACM Press.
- [OKI] OKI. Oki c 5900 color laser printer. <http://www.okiprintingsolutions.com>.
- [OS97] Kenton O'Hara and Abigail Sellen. A comparison of reading paper and on-line documents. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 335–342, New York, NY, USA, 1997. ACM Press.
- [Oxf] Oxford Easybook M3 Digital Notepad. Hamelin digital Conception Pragmatic. <http://www.oxfordeasybook.com>.
- [Pau89] Walter Pauk. *How to Study in College*. Houghton Mifflin, 1989.
- [Péd03] Roger T. Pédaque. Document: Form, sign and medium, as reformulated for electronic documents, 2003.
- [Peg] Pegasus Tablet NoteTaker. Pegasus Technologies Ltd. <http://www.pegatech.com>.
- [PKB05] I. Poupyrev, H. Kato, and M. Billinghurst. ARToolkit user manual, version 2, 2005.
- [Pla] Planon DocuPen RC 805. Planon System Solutions Inc. <http://planon.com>.
- [POK05] Annie Piolat, Thierry Olive, and Ronald T. Kellogg. Cognitive effort during note taking. *Applied Cognitive Psychology*, 19:291–312, 2005.
- [PS00] Réjean Plamondon and Sargur N. Srihari. On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):63–84, 2000.
- [RE94] Yvonne Rogers and Judi Ellis. Distributed cognition: an alternative framework for analysing and explaining collaborative working. *Journal of Information Technology*, 9:119–128, 1994.
- [Rek97] Jun Rekimoto. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 31–39, New York, NY, USA, 1997. ACM.
- [RSW<sup>+</sup>97] Peter Robinson, Dan Sheppard, Richard Watts, Robert Harding, and Steve Lay. Animated paper documents. In *Proceedings of the 7th International Conference on human-computer interaction (HCI '97)*, pages 655–658, 1997.

- [SAS01] Lisa Stifelman, Barry Arons, and Chris Schmandt. The audio notebook: paper and pen interaction with structured speech. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 182–189, New York, NY, USA, 2001. ACM Press.
- [SB08] Jürgen Steimle and Oliver Brdiczka. Paper-centric interaction concepts for collaborative learning. In *Conference Proceedings of Mensch und Computer 2008*, pages 207–216, 2008.
- [SBM08a] Jürgen Steimle, Oliver Brdiczka, and Max Mühlhäuser. Coscribe: Using paper for collaborative annotations in lectures. In *The 8th IEEE International Conference on Advanced Learning Technologies (ICALT 08)*, pages 306–310. IEEE Press, New York, NY, 2008.
- [SBM08b] Jürgen Steimle, Oliver Brdiczka, and Max Mühlhäuser. Cross-media linking and tagging support for learning groups. In *Proceedings of Tenth IEEE International Symposium on Multimedia (ISM 2008) (Third IEEE Workshop on Multimedia Technologies for E-Learning (MTEL 2008))*, pages 714–719. IEEE Press, New York, NY, 2008.
- [SBM08c] Jürgen Steimle, Oliver Brdiczka, and Max Mühlhäuser. Digital paper bookmarks: Collaborative structuring, indexing and tagging of paper documents. In *CHI: ACM Conference on Human Factors in Computing Systems (CHI 2008) Extended Abstracts*, pages 2895–2900. ACM Press, 2008.
- [SBM08d] Jürgen Steimle, Oliver Brdiczka, and Max Mühlhäuser. Paper-centric structuring in learning processes. In *The 8th IEEE International Conference on Advanced Learning Technologies (ICALT 08)*, pages 685–687. IEEE Press, New York, NY, 2008.
- [SBM09a] Jürgen Steimle, Oliver Brdiczka, and Max Mühlhäuser. Collaborative paper-based annotation of lecture slides. *Journal of Educational Technology & Society*, 12(4):(to appear), 2009.
- [SBM09b] Jürgen Steimle, Oliver Brdiczka, and Max Mühlhäuser. CoScribe: Integrating Paper and Digital Documents for Collaborative Knowledge Work. *IEEE Transactions on Learning Technologies*, 2(3):174–188, 2009.
- [Sea69] John Searle. *Speech Acts*. Cambridge University Press, 1969.
- [Sem] Semapedia. <http://www.semapedia.org>.
- [Ses04] Werner Sesink. *In-formatio. Die Einbildung des Computers*. LIT Verlag, 2004.
- [SGAL07] Hyunyoung Song, François Guimbretière, Michael A. Ambrose, and Carl Lostritto. Cubeexplorer: An evaluation of interaction techniques

- in architectural education. In *Human-Computer Interaction INTERACT 2007*, 2007.
- [SGF<sup>+</sup>09] Hyunyoung Song, Tovi Grossman, George Fitzmaurice, François Guimbretière, Azam Khan, Ramtin Attar, and Gordon Kurtenbach. Penlight: Combining a mobile projector and a digital pen for dynamic visual overlay. In *ACM Conference on Human Factors in Computing (CHI'2009)*, 2009.
- [SGHL06] Hyunyoung Song, François Guimbretière, Chang Hu, and Hod Lipson. Modelcraft: capturing freehand annotations and edits on physical 3d models. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 13–22, New York, NY, USA, 2006. ACM Press.
- [SGM07] Jürgen Steimle, Iryna Gurevych, and Max Mühlhäuser. Notetaking in University Courses and its Implications for eLearning Systems. In Christian Eibl, Johannes Magenheimer, Sigrid Schubert, and Martin Wessner, editors, *DeLFI 2007: 5. e-Learning Fachtagung Informatik*, pages 45–56, 2007.
- [SH03] Abigail J. Sellen and Richard H.R. Harper. *The Myth of the Paperless Office*. MIT Press, Cambridge, MA, USA, 2003.
- [SHM09] Daniel Schreiber, Melanie Hartmann, and Max Mühlhäuser. MundoMonkey: Customizing Interaction with Web Applications in Interactive Spaces. In *ACM SIGCHI conference Engineering Interactive Computing Systems (EICS)*, <http://www.acm.org/>, 2009. ACM Press.
- [Shn83] B. Shneiderman. Direct manipulation: A step beyond programming languages. *Computer*, 16(8):57–69, 1983.
- [Sig06] Beat Signer. *Fundamental Concepts for Interactive Paper and Cross-Media Information Spaces*. PhD thesis, ETH Zurich, 2006.
- [SJH<sup>+</sup>08] Thomas Seifried, Matthew Jervis, Michael Haller, Masood Masoodian, and Nicolas Villar. Integration of virtual and real document organization. In *TEI '08: Proceedings of the 2nd international conference on Tangible and embedded interaction*, pages 81–88, New York, NY, USA, 2008. ACM.
- [SLCGJ04] Orit Shaer, Nancy Leland, Eduardo H. Calvillo-Gamez, and Robert J. K. Jacob. The tac paradigm: specifying tangible user interfaces. *Personal Ubiquitous Comput.*, 8(5):359–369, 2004.
- [SMF99] Itiro Siio, Toshiyuki Masui, and Kentaro Fukuchi. Real-world interaction using the fieldmouse. In *UIST '99: Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 113–119, New York, NY, USA, 1999. ACM.

- [SN07] Beat Signer and Moira C. Norrie. Paperpoint: A paper-based presentation and interactive paper prototyping tool. In *Proceedings of TEI 2007, First International Conference on Tangible and Embedded Interaction*, pages 57–64, 2007.
- [Ste09] Jürgen Steimle. Designing pen-and-paper user interfaces for interaction with documents. In *Proceedings of the Third International Conference on Tangible and Embedded Interaction (TEI '09)*, pages 197–204. ACM Press, 2009.
- [Tay08] Stuart Taylor. Txt-it notes: Paper based text messaging. Technical Report MSR-TR-2008-172, Microsoft Research, 2008.
- [TBK04] Mrugank V. Thakor, Wendy Borsuk, and Maria Kalamas. Hotlists and web browsing behavior - an empirical investigation. *Journal of Business Research*, 57:776–786, 2004.
- [TLM09] Theophanis Tsandilas, Catherine Letondal, and Wendy E. Mackay. Musink: Composing music through augmented drawing. In *ACM Conference on Human Factors in Computing (CHI'2009)*, pages 819–828, 2009.
- [UI99a] Brygg Ullmer and Hiroshi Ishii. mediablocks: tangible interfaces for online media. In *CHI '99: CHI '99 extended abstracts on Human factors in computing systems*, pages 31–32, New York, NY, USA, 1999. ACM.
- [UI99b] John Underkoffler and Hiroshi Ishii. Urp: a luminous-tangible workbench for urban planning and design. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 386–393, New York, NY, USA, 1999. ACM Press.
- [UI01] Brygg Ullmer and Hiroshi Ishii. Emerging frameworks for tangible user interfaces. In John M. Carroll, editor, *Human-Computer Interaction in the New Millenium*. Addison-Wesley, 2001.
- [vD97] Andries van Dam. Post-wimp user interfaces. *Communications of the ACM*, 40(2):63–67, 1997.
- [VN00] F. Vernier and L. Nigay. A framework for the combination and characterization of output modalities. In *Proceedings of DSV-IS2000*, pages 32–48. Springer, 2000.
- [VV06] Melanie Volkamer and Roland Vogt. New generation of voting machines in germany – the hamburg way to verify correctness. In *Proceedings of the Frontiers in Electronic Elections (FEE 2006) Workshop*, 2006.
- [Wac] Wacom Graphire. <http://www.wacom.com>.
- [Wel93] Pierre Wellner. Interacting with paper on the digitaldesk. *Communications of the ACM*, 36(7):87–96, 1993.

- [Win86] Terry Winograd. A language/action perspective on the design of cooperative work. In *CSCW '86: Proceedings of the 1986 ACM conference on Computer-supported cooperative work*, pages 203–220, New York, NY, USA, 1986. ACM Press.
- [WISN08] Nadir Weibel, Adriana Ispas, Beat Signer, and Moira C. Norrie. Paper-proof: a paper-digital proof-editing system. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 2349–2354, New York, NY, USA, 2008. ACM.
- [Wiz] Wizcom InfoScan2 Elite pen. WizCom Technologies Ltd. <http://www.wizcomtech.com>.
- [WNS07] Nadir Weibel, Moira C. Norrie, and Beat Signer. A model for mapping between printed and digital document instances. In *Proceedings of DocEng 2007, ACM Symposium on Document Engineering*, 2007.
- [Wol00] Joanna L. Wolfe. Effects of annotations on student readers and writers. In *DL '00: Proceedings of the fifth ACM conference on Digital libraries*, pages 19–26, New York, NY, USA, 2000. ACM Press.
- [WQK07] David West, Aaron Quigley, and Judy Kay. Memento: a digital-physical scrapbook for memory sharing. *Personal Ubiquitous Computing*, 11:313–328, 2007.
- [WSS97] Lynn D. Wilcox, Bill N. Schilit, and Nitin Sawhney. Dynamite: a dynamically organized ink and audio notebook. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 186–193, New York, NY, USA, 1997. ACM Press.
- [YBK<sup>+</sup>06] Ron B. Yeh, Joel Brandt, Scott R. Klemmer, Jonas Boli, Eric Su, and Andreas Paepcke. Interactive gigapixel prints: Large paper interfaces for visual context, mobility, and collaboration. Technical report, Stanford University HCI Group, 2006.
- [YLK<sup>+</sup>06] Ron Yeh, Chunyuan Liao, Scott Klemmer, François Guimbretière, Brian Lee, Boyko Kakaradov, Jeannie Stamberger, and Andreas Paepcke. Butterflynet: a mobile capture and access system for field biology research. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 571–580, New York, NY, USA, 2006. ACM Press.
- [YPK08] Ron B. Yeh, Andreas Paepcke, and Scott R. Klemmer. Iterative design and evaluation of an event architecture for pen-and-paper interfaces. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 111–120, New York, NY, USA, 2008. ACM.

# Index

- A-Book, 66, 74, 80, 83
- Abstracting, 15, 48, 150, 161
- Acceptance, 192, 202
- Active reading, 3, 48, 132, 159
- Adaptivity, 134
- Affordance, 16, 18, 43, 109, 132
- Air traffic control, 18
- Annotation, 2, 3, 15, 17, 19, 25, 29, 30, 32, 44, 80, 92, 130, 194, 212
  - clustering, 30, 186, **187**
  - handwritten, **130**, 132, 186
  - style, 31
- Anoto pattern, **61**, 63, 121, 162, 181
- Anoto pen, 61, **62**, 63, 104, 113, 120, 139, 184, 195
- Archiving, 2, 19, 25
- Arrangement, 17, 37, 44, 124, 213
- ARTag, 56, 186
- ARToolkit, 56
- Association area, 152, **155**, 170
- Association gesture, 152, 170
- Audio, 2, 43, 63, 65, 67, 72, 82, 100, 104, 123, 155
- Audio Notebook, 65
  
- Barcode, 55, 70, 71, 156
- Book, 2, 71, 100, 114, 155, 158
- Bookmark, 44, 118, 129, 162
- Books with Voices, 71
- Bush, Vannevar, 6, 44, 73, 111, 151
- ButterflyNet, 65, 95
  
- CARE properties, 103
- Channel, 89, 110
- Character error rate, 145
- Client/server system, 181
- Cognitive load, 46, 84, 191, 193
- Cohabitation, 5
- Cohen's Kappa, 189
- Collaboration, 6, 15, 18, 19, 28, 42, 45, 83, 105, 116, 135, 159, 169, 199, 212, 217
  - co-located, 3, 34, 105, 171
  - remote, 3, 107, 177
  - sharing, *see* Sharing
- Complementarity of information, 102
- Conceptual activity, 43, 80, 91, **92**, 95
- Context, 2, 19, 32, 46, 113
- Core interaction, 92, **93**, 95, 96, 109, 217
  - altering shape, 93, 94
  - bridging, 94, 158, 164, 171
  - clicking, 93
  - combining, 94, 158, 164
  - inking, 93, 158, 164, 171
  - moving, 93, 94
- Cornell Notetaking Method, 17
- CoScribe, 6, 87, 92, 100, 104, **111**, 194, 212
- CoScribe print toolkit, 114, 133, 181
- CoScribe viewer, 113, 123, 134, 140, 155, 158, 183
- Cost factor, 19

- Cube, *see* Process cube
- DataMatrix, 56
- Design guidelines, 109
- Designer's Outpost, 68
- Dictionary, 147
- Digital Desk, 59, 67, 72, 83, 95
- Digital Paper Bookmark, 128, **162**, 183, 201
- Digital pen, *see also* Anoto pen, 59
- Direct manipulation, 88, 179
- Discontinuity
  - of media, 99, 100
  - of tools, 120
- Display
  - Anoto-enabled, 123, 184, 203, 206
  - Tabletop
    - Occlusion, 215
- Distributed Cognition, 39
- Document, 34, **99**, 121, 155
  - digital-only, 100
  - function, 35
  - hybrid, 100, 124, 155, 156
  - ownership, 37
  - physical-only, 100, 156
  - printed-only, 125
  - space, 3, 35, 161
- E-Book reader, 17, 42, 120
- EAN, 55
- Ease of learning, 193
- Ease of use, 193
- Eclipse, 36
- Ecological perspective, 4, 38, **41**, 177, 212
- Ecological view, **124**, 127, 158, 183
- Embodiment, 16
- Encoding (cogn.), 44
- ePaper, 216
- Equivalence of information, 101
- Externalization, 35
- Feedback, 43, 109, 155
- Field study, 4, 7, 19, 41, 150, 211
- Filter, 126
- Fly Fusion Pentop Computer, 63
- Folder, 44, 118, 125, 128, 155, 157, 158, 183
- Functional system, 39
- Gender, 22
- Graphical User Interface, 52, 88, 90, 94, 109, 134, 135, 179
- Greasemonkey, 183
- Grounded theory, 34
- Group work, 28, 34
- Guidance, 14
- Handwriting recognition, 110, **144**, 155, 170, 187
- Haptic, 43, 103, 104
- Highlighting, 3, 17, 30, 32
- Hyperlink, 3, 44, 69, 75, 76, 80, 92, 125, 150, **151**, 203, 205, 212, 214
- Hypertext, 151
- Indication
  - map-based, 128
  - property-based, 128
- Information ecology, **40**, 42, 46, 84, 111, 161, 177
- Information retrieval, 203
- Instructor, 21, 133, 144
- Instructor interface, 144
- Instrumental Interaction, 96
- Interaction
  - semantic level, 91, 96, 109
  - syntactic level, 91, 96, 109
- Interaction model, 88, 96, 212
- Interaction primitive, 92
- Interaction technique, 5, 88
- Interview, 31, 34, 194, 199, 204
- Java, 181, 183
- JavaScript, 183
- Keyboard, 31, 83, 106, 195, 198, 215
- Knowledge work, 1, 2, 4, 13, **14**, 16, 113
  - paper-based, **88**

- Language, 27  
Laptop, 19, 23, 24, 30, 35, 212  
Layout, 18, 134, 213  
Learner, 21  
Learning, 14, 115, 150, 161, 212  
    forms, 14  
    group meeting, 34, 115, 116  
    quality, 192  
Lecture, 20, 193, 199  
Levenshtein Distance, 145  
Livescribe, 63, 65  
Logical page, 181, 186
- Marble Answering Machine, 53  
MediaBlocks, 53  
Memex, 6, 73  
Meta-cognition, 162, 166, 172  
Mobility, 18, 83, 134, 214  
Model-View-Controller, 52, 88, 183  
Mouse, 83, 106, 121, 184, 198, 215  
Mozilla Firefox, 123, 156, 158, 183  
Multi-user view, 140, 199  
Multiplexing, 97, 109, 158  
MundoCore, 183
- Navigation, 2, 13, 16, 47, 101, 126, 159,  
    167, 199  
Nelson, Ted, 73, 111, 151  
Note, 3, 17, 25, 64  
Notetaking, 15, 19, 44, 64, 80
- Optical character recognition (OCR),  
    57, 75  
Organon model, 21
- Page area, 94  
Paper Augmented Digital Documents,  
    67  
Paperless office, 1, 15  
PaperPDA, 75, 80  
PaperPoint, 69  
PapierCraft, 74, 77, 81, 83, 84, 95  
PDA, 43, 66  
PDF, 84, 100, 114, 183
- Pen-and-Paper User Interface, 5, 8, 43,  
    51, 88, 135, 212  
Perceptual coupling, 82  
Physical page, 181  
PowerPoint, 20, 29, 84, 114, 123, 183  
PPUI, *see* Pen-and-Paper User Interface  
Print toolkit, *see* CoScribe print toolkit  
Printer, 183  
Privacy, 84, 108, 110  
    visibility, 137, 198  
Process cube, 173  
Process knob, 175  
Process structuring, 35, 172
- Quality, *see* Learning  
Questionnaire, 20, 194, 204  
Quickies, 78
- Reading, 3, 15–17  
Reference, 35, 92, 151  
Relation, 42, 44, 92  
Representation, 52, 99, 101–103, 134  
Requirement, 7, 42, 119, 131, 151, 163,  
    191  
Review, 27, 134, 199  
RFID, 56, 70, 72, 173
- Satisfaction, 192  
Scalability, 143  
Scanner, 57  
Scope, 92, 155  
Search, 2, 45, 123, 147, 161  
    full-text, 19, 126, 144, 149  
Semantic category, 30, 32, 162, 172, 174,  
    197  
Seminar, 29  
Sharing, 2, 14, 108, 135, 139, 159, 162,  
    166, 169, 171, 172, 177, 213  
Single-user view, 140, 199  
Sticker, 44, 75, 118, 128, 162  
Storage (cogn.), 44  
Structuring, 2, 3, 42, 44, 47, 118, 133,  
    161, 162, 169  
    process, *see* Process structuring  
Symbol, 32, 44

- Tablet PC, 2, 17, 24, 42, 67, 83, 120, 187
- Tabletop display, *see* Display
- TAC paradigm, 53
- Tag, 125, 126, **162**
- Tag Button, **138**, 197, 198
- Tag Menu Card, 122, 135, **170**, 183
- Tagging, 3, 44, 76, 81, 92, 161, 197, 212
  - category, 162
  - free, 162
- Tangible tool, 114, 121, 173
- Tangible User Interface, 4, 52, 88
- TUI, *see* Tangible User Interface
  
- Ubiquitous Computing, 4, 88
- Unified pen-based interaction, 6, 43, 83, **118**, 206
- URP, 53
- User study, 6, 191
- User-centered design, 6
  
- Video, 2, 71, 100, 123
- Viewer, *see* CoScribe viewer
- Visualization, 5, 118, 167, 172, 213
  
- Web 2.0, 77, 217
- Web page, 100, 114, 123, 183, 203
- WebStickers, 70
- Wikipedia, 35, 204
- WIMP model, 88
- Word error rate, 145

# Bildungsgang des Autors

Jürgen Steimle, geboren in Mutlangen

Studium der Informatik und der Romanistik an der Albert-Ludwigs-Universität Freiburg und der Université Lyon 2, Lyon, Frankreich. Stipendiat der Studienstiftung des Deutschen Volkes.

Vorbereitung der Promotion unter Betreuung von Prof. Dr. Max Mühlhäuser als Wissenschaftlicher Mitarbeiter am Fachbereich Informatik der Technischen Universität Darmstadt. Stipendiat im DFG Graduiertenkolleg E-Learning.