

Cross-Media Linking and Tagging Support for Learning Groups

Jürgen Steimle¹, Oliver Brdiczka², Max Mühlhäuser¹

¹*Technische Universität Darmstadt*
 {steimle, max}@tk.informatik.tu-
 darmstadt.de

²*Palo Alto Research Center (PARC)*
 oliver@parc.com

Abstract

Typical tasks of learning groups and knowledge workers include working both with printed and digital documents. We present a pen-based and tangible interaction concept for the linking and tagging of documents in a mixed physical and digital environment. The interaction with printed and electronic documents is unified. We therefore use digital pens, which equally capture physical handwriting on real paper and act as an input device on a specific screen prototype. Links and tags can have variable scopes ranging from small document passages to a collection of several documents. Physical folders along with camera-based marker tracking provide for the tangible definition of document collections. Tags are intuitively defined on a paper tag menu card. We present two visualizations of links and tags (a document-centered viewer and a hyperstructure-centered graph view). These are closely coupled with the physical environment.

1. Introduction

An ethnographic study [14] of learning group meetings found that meetings typically rely on a considerable number of both printed and digital documents, which are strongly interrelated with each other (Fig. 1 left). This includes for example printed or digital lecture scripts and articles, physical textbooks, digital PDF and Web documents on one or several laptops, and empty sheets of paper for taking notes. These printed and digital documents form one document space covering the problem domain.

Structuring this problem domain, abstracting, and establishing relationships between concepts is crucial for successful learning. The resulting structural knowledge facilitates recall and comprehension and is essential to problem solving [7]. 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.

So far, practices and tools for the support of structuring and relating documents are restricted either to paper documents or to digital documents and lead to

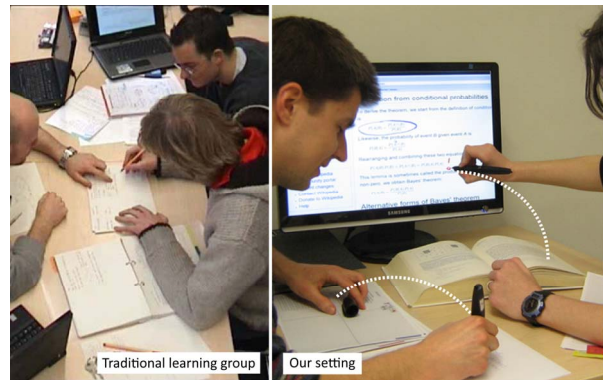


Figure 1: Our setting is inspired by media use in typical learning group meetings (left). Associative pen gestures provide for linking and tagging in a mixed physical and digital environment (right).

artificial gaps between media. Moreover, activities with physical documents are bound to their physical instance and therefore constrain collaboration.

In this paper, we present a pen-based and tangible interaction design that unifies both worlds in a mixed physical and digital environment. Linking and tagging is made with a digital pen. This pen is the central interaction device that offers the same interactions for working with physical documents on traditional paper and with digital documents on a specific screen. The focus lies now on interacting with *documents*, regardless of their representation. The hybrid physical and digital hyperstructure created by all users of one learning group is electronically integrated in digital document views and a graph visualization. This supports the efficient paper-based and screen-based browsing of the collaborative hyperstructure.

2. Related Work

Interaction Theory. The theoretical basis for our interaction design is the concept of Information Ecologies [11] and the Distributed Cognition framework [6]. Roughly speaking, these state that in a given collaborative work context, physical and digital artifacts, people, and practices are strongly interrelated. As even minor changes of practices can have major and often unforeseen consequences for an established work practice, ethnographic studies are an important part of our work.

Our own and other research [10] show that the way how people work with documents, e.g. annotating and tagging, is highly individual. The key principle of our design is therefore to depart from existing practices and to leave much freedom to the users. However, a slight formalization of some interactions is necessary to enable computers to interpret them. We therefore identify appropriate minor formalizations in the user interface design that are acceptable and can be symbiotically integrated into the ecology.

Cross-media hyperlinks and tags. There has been substantial research on combining hyperlinking and tagging with the tangible manipulation of *real physical documents*. Contrary to our system, most work focuses on a single document type and on links from paper to digital documents only, but does not support the broader document space at typical workplaces. Books with Voices [8] lets users access audio files from pre-defined link hot-spots on a paper document. PaperLink [2] relies on an enhanced pen and supports creating and browsing links from printed to digital media, but not vice versa. PapierCraft [9] presents a pen-based command interface for copy/pasting, linking, and tagging documents, but remains limited to paper. A prototype for Physical Hypermedia [5] supports users in linking and grouping physical objects by holding them close to an RFID reader. As the focus lies on general physical objects, a more specific document support (e.g. linking sub-passages) is missing. The iServer framework [12] presents an extensive model for cross-media links between physical and digital documents, but remains silent on interactions for creating hyperlinks.

Unified interaction on paper and screens. In contrast to our approach, all systems mentioned require separate interactions and devices for paper and digital documents. An exception is Wellner's Digital Desk [16], which overlays printed documents with digital information and provides pointing gestures as a unified interaction for printed and digital documents. To the best of our knowledge, Voodosketch [3] is the first system that uses digital Anoto™ pens as an integrated interaction device both for a screen and for paper palettes. However, the interaction remains centered to digital documents, as the paper palettes only contain additional controls. In contrast, our design supports physical and digital documents at an equal footage.

3. System Design

The design of our research prototype was guided by the following goals:

Unified and intuitive interaction: Working with paper and digital documents shall rely on the same interaction technique and device, as the user's focus lies on working with documents, not with devices. The interaction should be intuitive and easy to integrate into current work practices.

Annotation, tagging, and cross-media hyperlinks: The system shall enable users to annotate and tag documents and to structure document sets by creating and browsing hyperlinks. This applies to both physical and digital documents.

Collaboration: Annotations, tags, and cross-media hyperlinks shall be shareable. This enables users to digitally share their structuring of both physical and digital documents over the distance.

Our design supports typical learning group settings but also the setting of individual knowledge work. One or several persons work at a table or a desk both with printed and digital documents. Physical documents are located on the desk. Digital documents can be displayed on a shared screen (Fig. 1 right) or tabletop display. Each user has its own digital pen for annotating, linking, and tagging documents. This acts both as ball pen on paper and as an input device on the display. The core functionality for interacting with physical documents does not require the shared display or a nearby computer and hence supports mobile use.

The system is based on an earlier prototype, which allows to collaboratively annotate printed documents [15]. This is now extended by tags and hyperlinks. A software viewer provides access to a digital copy of the paper documents along with both own and shared annotations, tags, and hyperlinks (Fig. 6). This viewer supports the same pen-based interactions that apply to printed documents. Moreover, tags and links from, to and within Web pages can be established within Mozilla Firefox on a tabletop display. We therefore developed a Firefox plug-in that manages and displays pen-based hyperlinks. Plug-ins for other document types can be easily integrated into the architecture.

3.1. A Unique Interaction Device

Our approach unifies working with printed and digital documents: one digital Anoto™ pen acts both as a ball pen and e-Ink capturing device on paper printouts and as an input device on a display (Fig. 1). Contrary to previous research, users do not need to switch between different input devices when working with printed and digital documents.

The Anoto technology for pen position detection on printouts relies on a specific dot pattern printed on paper. The pen identifies both the page and its position on this page by detecting a specific, nearly invisible dot pattern with a built-in camera. The connection to a nearby computer is made via Bluetooth and streams the stroke data in real-time. Alternatively, a USB connection can be used. In contrast to other solutions, users can naturally work with multiple paper sheets like with traditional pen and paper. This property is crucial for linking different documents.

Following the approach of [3], we constructed a display, on which the same pen can be used as an input device. This was realized by printing the dot pattern on a translucent film that is illuminated by rear-projection. As pen data originating from the screen is translated to mouse events, not only our own system can be controlled with the pen, but also all other applications.

3.2. Model of Links and Tags

Our motivating design model is to conceptualize both links and tags as undirected binary associations. A link associates two document scopes (which is a part of a document, an entire document, or a collection of documents) with each other. It is visualized on both endings. By tipping on this marking, the target document is displayed. A tag associates a document scope with a representation of a concept. In our case, this is a label of one or several keywords.

Associations are created by performing a pen gesture that connects two association areas. The types of these areas decide on the type of the association. Links are created by connecting two document association areas. Tags are created by connecting a document association area with a representation of the tag. Different types of document association areas further provide for defining the document scope the link or tag applies to.

In order to integrate the various different media of a typical learners' document space, the interaction design supports associations between a wide range of document types. Currently these are a) *digital-only documents* without physical counterpart (Web pages) b) *physical-only documents*, which are digitally represented by some meta-data (books, journals) and c) *combined physical and digital documents* with nearly identical representations in both worlds, whose printed instances are covered by Anoto pattern (PDF and PPT documents and empty sheets of paper). To support interactions with Anoto pens on books, users can attach a small sticker covered with Anoto pattern onto the book cover and register this with the barcode number of the book. In the next two sections, we will first

present the pen-based association gesture (*operator*) and then detail on the association areas (*operands*), where this gesture is made.

4. A Pen-based Association Gesture For Linking and Tagging

Accounting for the associative nature of links and tags in our model, the central interaction for creating links and tags is a pen-based association gesture. This connects two association areas (on paper or on a screen) (Fig. 1 right). These areas are described in the next section. Our interaction model considers both a paper page and a screen as similar display instances. It is thus an extension to prior "stitching" gestures that span different screens [13] or documents [9].

In order to create a hyperlink, the pen performs one of the following gestures: If the documents are located close to each other, they can be linked with a single straight line (Fig. 2 left). Alternatively, one can split up the gesture into two consecutive point markings on each of the two areas (Fig. 2 right). The link gestures are recognized in real-time and instant audio feedback is given. This is particularly important when working with paper documents, as no visual feedback can be provided. Links can be deleted with a cross-out gesture on either ending of the association.

A hyperlink is activated by tipping on the marking in an association area (point or line). This is possible both on a printed document page and on a digital document on the screen. The target document is then displayed either in the web browser or in the software viewer for the digital representation of paper documents. The document which the hyperlink refers to is highlighted in yellow. Backward and forward buttons support navigating in a cross-media history that contains not only web pages, but also pages of PDF and PPT documents displayed in the software viewer.

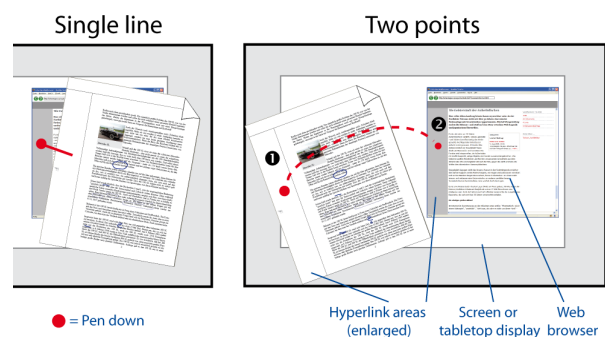


Figure 2. Pen gestures for creating an association (red) are possible on any combination of printed and digital documents.

5. Association Areas

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 margins of the document, leading to clearly visible references.

Inspired by this practice, we opted for a spatial differentiation between association gestures and the actual document. A key issue of our interaction design is thus the separation between a generic gesture and different association areas, which act as end point for the gesture and decide upon the parameters of the association. This combines intuitive interaction with versatility and predictability even with the restricted feedback capabilities of a paper-based user interface.

In the remainder of this section, we will define first the document association areas and second the association area for creating tags.

5.1. Document Association Areas

Different association areas on documents represent different document granularity levels. Taking the document as the unit of reference, our model distinguishes three scope levels: A scope can be an entire document, a subpassage of a document, or a collection of documents:

Document level. Each document has an association area on top of its first page. Association gestures made within this area refer to the entire document. Both printed and digital documents include this area (Fig. 3 2nd line). Another digital association area is the document thumbnail in the graph visualization of links and tags (see next section).

Sub-document level. The association area that represents the sub-document level is a small column located to the left or to the right of the actual document (or on both sides in case of two-column documents) (Fig. 3 3rd line). Association gestures that are made in this area refer to a contiguous passage within a document. An optional vertical marking near the association marking defines the precise extent of this passage.

Super-document level. This granularity level is a collection of multiple documents. A very intuitive traditional interaction for defining a collection of paper documents is to physically co-locate them in a folder or binder. Users can then easily reference the entire collection by referring to the folder.

Our slightly formalized interaction design is inspired by this co-location. Several printed documents can be put into one paper folder to define a collection (Fig. 4a). The system automatically detects this activity. This is done by using a camera, which is located

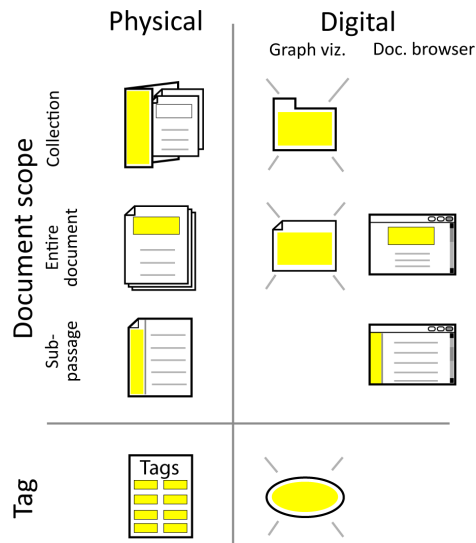


Figure 3. Association areas for pen gestures (highlighted in yellow)

above the desk and which tracks 2D barcodes that are printed on the cover of each folder and on the first page of each document. Paper folders have various colors and can optionally be labeled with a handwritten title. Associations with the collection are then made by performing the association gesture on the association area on the front flap of the folder (Fig. 3 top). An alternative association area is the digital representation of a folder in our graph visualization. This way, folders can be tagged and linked like ordinary documents.

5.2. Tag Menu Card

Tags are created by associating a document scope with a tag association area. A separate tag menu card (Fig. 4) represents this area. Users can define a new tag at any time by writing a keyword on this paper card. Depending on the setting, each user can have one or more own cards or multiple users can share a card. Tagging is performed by performing one part of the association gesture on the appropriate tag keyword (Fig. 3 bottom). An alternative digital association area is the representation of a tag in the graph visualization (see next section). To support collaborative tagging with individual tag menu cards, tag keywords of multiple users can be declared as identical by performing an association gesture on these keywords.

In contrast to writing a tag keyword directly into the document, collecting all tags on a separate tag menu card has several advantages. First, the user can immediately access a set of all her tags. Second, research shows that a key factor for the convergence of tags is that the system suggests frequent labels [4]. As it is faster to associate an already existing tag than defining

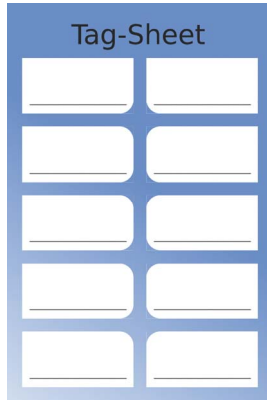


Figure 4: Tag Menu Card

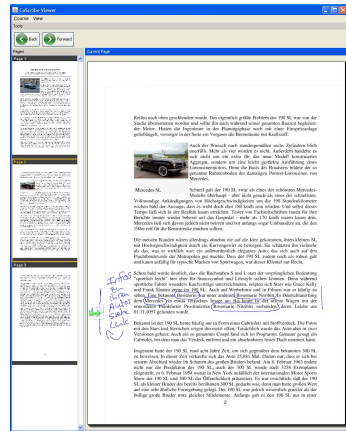


Figure 5: Paper document viewer.

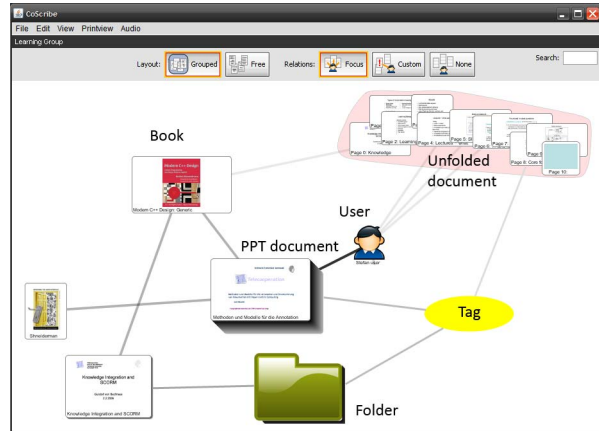


Figure 6: Graph visualization of the hyperstructure.

a new one, we assume that a tag menu card can have similar effects as the automatic suggestion, which is not possible in a pure paper user interface.

6. Visualization of the Hyperstructure

All linking and tagging activities are automatically integrated into two visualizations that can be displayed on the shared screen. Both visualizations are accessible from paper documents by tipping with the pen on a corresponding button, which is printed on each page.

The first visualization focuses on working with an individual document. Individual documents are displayed in a viewer for paper documents along with handwritten annotations of all users, links, and tags (Fig. 5). Web pages are accessible in Mozilla Firefox.

The second visualization is a graph view that focuses on the entire document collection (Fig. 6). Nodes are thumbnail representations of documents and document pages as well as representation of folders, tags and users. Edges correspond to the associations between those node instances. Edges between documents express links, edges between a document and a tag node express a tag, and edges between a document and a user node express that this user has worked on this document by adding annotations, tags or hyperlinks. The node for a printed document can be unfolded to display the individual pages. Similarly, the individual documents within a folder can be accessed. Direct manipulation provides for flexible panning and zooming as well as for centering or repositioning of individual items. Double clicking on document nodes displays the viewer or Web browser with this document. Moreover, comprehensive filter options are available. Both visualizations support collaboration and the co-construction of structure, as all links and tags of all users are integrated into one view.

7. Implementation and Early User Feedback

We have implemented a working prototype. The system is developed in Java. We use the Anoto SDK and modules of PaperToolkit¹ for the synchronization of pen strokes. Pen gestures are detected by heuristics. The graph visualization was implemented with the Prefuse framework¹. All data is stored in a graph-based database. We perform the marker-based tracking of document positions with ARTag¹. The Mozilla Firefox plug-in was realized in JavaScript. The communication with the Java application is based on the MundoCore middleware [1].

During design and implementation, we gathered early user feedback. Paper mockups and working modules for the pen-based association and the visualization were reviewed with 11 computer science, pedagogy and law students and researchers of our institution. All but one participant found it helpful to have separate annotation and association areas. They valued the more intuitive and clear interaction in contrast to mode switching that would be necessary otherwise. Moreover, association areas help to quickly find positions of links and tags in the printed document.

The participants found it important to be able to define scopes of variable sizes. All participants from computer science and pedagogy judged scopes down to the level of individual paragraphs as sufficient. However, a law student needed granularities down to word level. For this student, it was acceptable to define the corresponding line as computer interpretable scope and to make an additional uninterpreted annotation that highlights the actual word. Two participants don't wish computer interpretation of their scopes at all. For them,

¹ hci.stanford.edu/paper; www.artag.net; prefuse.org

fuzzy scopes that they can interpret themselves are sufficient. In contrast to the association area for an entire document, which is situated on the first page, an association area for each single page was not judged necessary. This is due to the fact that a page typically is not a self-contained section.

The concept of tagging documents with the tag menu card was appreciated. An explicit functionality for writing tags directly onto a document was not considered missing because this can be simulated by writing down an arbitrary label and additionally perform an association to a tag on the tag menu card. Two users expressed the wish for a possibility to define tag sets. We will investigate this further.

8. Conclusion

In this paper, we presented novel interaction techniques for linking and tagging paper and digital documents. The interactions are based on a digital pen that writes both on paper and on a specific screen. Two visualizations integrate physical and digital documents and support users in browsing the cross-media hyperstructure. Our approach closely integrates printed and digital documents to a novel user experience, which leads us a major step towards closing the gap between paper and the digital.

9. Acknowledgments

The authors thank Michael Stieler and Daniel Schreiber for their help on implementing the prototype. We are also grateful to Christoph Koenig and Petra Gehring for helpful discussions. This work was funded by the German Research Foundation as part of the Graduate School on *Feedback-Based Quality Management in eLearning* (DFG-GK-1223).

10. References

[1] E. Aitenbichler, J. Kangasharju, and M. Mühlhäuser, "MundoCore: A Light-weight Infrastructure for Pervasive Computing," *Pervasive and Mobile Computing*, 2007.

[2] T. Arai, D. Aust, and S. 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*. New York, NY, USA: ACM, 1997, pp. 327–334.

[3] F. Block, M. Haller, H. Gellersen, C. Gutwin, and M. Billinghurst, "Voodoosketch: Physical interface palettes and sketched controls alongside augmented work surfaces," in *UbiComp 07*, 2007.

[4] S. A. Golder and B. A. Huberman, "The structure of collaborative tagging systems," *HP Labs, Tech. Rep.*, 2005.

[5] K. Grønbaek, J. F. Kristensen, P. Ørbæk, and M. A. Eriksen. "Physical hypermedia": organising collections of

mixed physical and digital material," in *HYPERTEXT '03: Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*. New York, NY, USA: ACM, 2003, pp. 10–19.

[6] J. Hollan, E. Hutchins, and D. Kirsh, "Distributed cognition: Toward a new foundation for human-computer interaction research," *ACM Transactions on Human-Computer Interaction*, vol. 7, no. 2, pp. 174–196, 2000.

[7] D. H. Jonassen, K. Beissner, and M. Yacci, *Structural Knowledge: Techniques for Representing, Conveying, and Acquiring Structural Knowledge*. Hillsdale: Lawrence Erlbaum Associates, 1993.

[8] S. R. Klemmer, J. Graham, G. J. Wolff, and J. 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*. New York, NY, USA: ACM Press, 2003, pp. 89–96.

[9] C. Liao, F. Guimbretière, K. Hinckley, and J. Hollan, "Papiercraft: A gesture-based command system for interactive paper," *ACM Transactions on Computer-Human Interaction*, vol. 14, no. 4, pp. 1–27, 2008.

[10] C. C. Marshall, "Toward an ecology of hypertext annotation," in *HYPERTEXT '98: Proceedings of the ninth ACM conference on Hypertext and hypermedia*. New York, NY, USA: ACM Press, 1998, pp. 40–49.

[11] B. Nardi and V. O'Day, *Information Ecologies: Using Technology with Heart*. MIT Press, 1999, Chapter Four: Information Ecologies.

[12] M. C. Norrie, B. Signer, and N. 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.

[13] J. 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*. New York, NY, USA: ACM, 1997, pp. 31–39.

[14] J. Steimle and O. Brdiczka, "Paper-centric interaction concepts for collaborative learning," in *Conference Proceedings of Mensch und Computer 2008*, 2008, pp. 207–216.

[15] J. Steimle, O. Brdiczka, and M. Mühlhäuser, "Co-scribe: Using paper for collaborative annotations in lectures," in *The 8th IEEE International Conference on Advanced Learning Technologies (ICALT'08)*. IEEE Press, New York, NY, 2008, pp. 306–310.

[16] P. Wellner, "Interacting with paper on the digitaldesk," vol. 36, no. 7. New York, NY, USA: ACM, 1993, pp. 87–96.