

# VirDO: A Virtual Workspace for Research Documents<sup>\*</sup>

George E. Raptis, Christina P. Katsini and Stephen J. Payne

Department of Computer Science, University of Bath, Bath, UK  
{george.e.raptis, christina.katsini}@bath.edu, s.j.payne@bath.ac.uk

**Abstract.** We report the design of a system which integrates a suite of tools to allow scholars to manage related documents in their personal digital stores. VirDO provides a virtual workspace in which pdfs can be placed and displayed, and which allows these documents to be manipulated in various ways that prior literature suggests to be useful. Particularly noteworthy are the various maps that support users in uncovering the inter-relations among documents in the workspace, including citation relations and exible user-dened tags. Early evaluation of the system was positive: especially promising was the increasing use of maps by two participants who used VirDO for their own research over a period of a week, as well as the extensive use by all participants of sticky notes.

**Keywords:** sensemaking, document mapping, annotation, scholarship

## 1 Introduction

Managing academic documents from the literature is a vital task for researchers everyday life. We present a system, VirDO (for Virtual Document Organizer) designed to help scientists manage a local store of downloaded documents, providing tools to support sensemaking.

Many tools have already been developed to support easier organization of digital documents [1,4]. However, only a few support sensemaking from local, individual document stores, which in the case of research is complicated by research literature having unique characteristics such as citations networks, particular categories of articles, pre-dened keywords etc. Scientists use some distinct processes in order to build their local research libraries, such as tracing citations, evaluating articles in terms of citation frequency, or relying on other peoples comments so they can decide whether a paper is worth reading or not.

Modjeska et al. [8] proposed ve features for an eective bibliographic visualization. Brievely, an eective bibliographic visualization tool should provide full bibliographic information, such as title and author(s); lter data; display chronology and citations of the articles; allow for detailed view of information supporting dierent levels of details and be capable of visualizing large search sets. These

---

<sup>\*</sup> The final publication is available at Springer via [http://dx.doi.org/10.1007/978-3-642-40501-3\\_63](http://dx.doi.org/10.1007/978-3-642-40501-3_63)

authors describe the BIVTECI system, which incorporated these features and influenced our own design, which nonetheless differs in several details.

As its name (Bibliographic Visualization Tool with enhanced Citation Interactivity) suggests BIVTECI attempts to support sensemaking through visualization of document spaces, with particular reference to citation links. A large number of systems have similarly explored these issues [4,8] including techniques such as growing polygons between related documents [4], family tree networks [7] and hypertext co-citation maps [3]. However, these tools have focused on computation and display of properties of very large document collections rather than limited local stores. The provision of an interactive workspace to replace a standard document folder is our focus.

In this respect VirDO is also similar to CiteSense [10], which is similar in intent to BIVTECI but additionally allows personal annotations, which facility we also provide through tags. File or document tagging has been widely studied recently, most often as an alternative retrieval mechanism to hierarchical folder structures [5]. In our system, tags provide an additional basis for mapping the document space, reflecting our focus on the organization of a local workspace. This is another issue that has seen a great deal of exploratory research, including studies of notes and real desks [6]. To very briefly summarize, this research has shown the importance of: (a) document location: Digital resources can be organized spatially too and location is used in retrieval [2]; (b) notes: people attach notes to paper documents and find this feature useful in electronic environments [9]; (c) filing and piling: people adapt their strategies to their task, and should be allowed to store documents in ad hoc piles which are better for some tasks [6]; (d) working versus archived documents: documents are important to their users over different timescales, from ephemeral to long-term [2,9]; (e) virtual workspaces: a promising idea is the creation of virtual workspaces, where the documents are organized in a more realistic manner [1].

## 2 VirDO

VirDO allows users to organize their electronic documents spatially, as they would with paper documents on their desks. It provides them with a virtual workspace, where they can perform various operations, such as re-arrangement of the documents, importing of new documents, deletion of documents they do not need any more, creation of spatial maps which show various relations between the documents and attachment of sticky notes. The virtual workspace consists of the following elements:

- Incoming Box: This is a place for all the documents of the project folder which the user has not yet placed on his workspace, because they are new and/or unclassified. The user can move these into the workspace either by dragging and dropping each of them, or by using the multiple import feature.
- Trash Bin: where the user can place whatever exists on his workspace and is not useful anymore; including digital documents, sticky notes, etc.

- Virtual Document: A virtual instance of a document can be placed into the workspace. Each is a 3-D box displaying a thumbnail of the first page of the actual document. Its length and height have fixed-size, but its depth is related to its number of pages, giving the user an immediate sense of document length. Moreover, each virtual document provides the user with more information about it. In particular two ways of representing such information are provided: a vertical bar indicating the number of the citations the documents has, and a popup menu offering a variety of choices
  - The vertical bar is updated regularly by web services, such as Google Scholar. The height of the bar indicates the number of citations as a proportion of the highest number of citations for any paper in the collection (which is fixed to the maximum-height bar).
  - Through the popup menu, the user can view the whole document with an integrated pdf viewer offering document navigation, markup, zoom, etc. The user can also view additional document properties: title, author(s), user-assigned tags, references and the papers that cite that document, along with the number of citations of each reference or citation-source. Moreover, the user can mark a document as read or not-read, post personal notes, or create any of four interrelation maps: by author(s), by citations, by tag(s) or by publication year.

Considering the maps, documents related by the chosen attribute are connected with lines showing the inter-relations. When Author is selected, that authors documents are connected; when Tags is enabled, the documents with that tag in common are connected; for Year, the documents published in the selected year are connected. Finally, Citations shows a selected document linked both with its references and its citations. The maps are created in the main workspace and in a new tab window of VirDO, where only the interrelated documents are displayed, re-arranged in the space.

An additional feature of VirDO is the creation and use of notes. The user can create sticky notes and place them anywhere in the workspace or even on the documents, as he would normally do with the paper documents, creating in that way piles, or spatially-close clusters of interrelated documents.

### 3 Evaluation study

For the first evaluation of VirDO we conducted two small empirical studies, a shorter and a longer term study, in which participants used the system to consider a small sub-part of an academic literature. Participants in both studies were given a brief training session, by watching a video-tutorial of the system in use, and all interactions with VirDO were recorded.

In the shorter-term study, thirteen graduate students had already each used Google Scholar to download a set of pdfs (5 to 35) on a topic provided by the investigators. They spent 20 minutes using VirDO to organize these pdfs. - we observed extensive use of tags and sticky notes by all participants. Only five participants used the maps, most commonly the Tags map. In the longer-term study,

two graduate students were invited to use VirDO to organize their preexisting locally stored set of academic documents, as well as any new documents they retrieved during the study. Their use of VirDO was recorded for a full week.

In the later stages of this study, both participants made use of maps, and their use gradually increased. Tag maps and Citation maps were used regularly, whereas Author and Year maps were used less often. The longer-term participants commented on the visualization of the document relationships. They found connecting lines useful and effective, but advocated colour-coding lines. Finally, opening a new tab when a new map was generated was found to be useful.

## 4 Conclusion

Our main aim in this work has been to explore the idea that the academic sensemaking process can fruitfully be considered in the context of relatively small local collections of academic documents. The iterative building of local document stores requires incremental, iterative sensemaking of document spaces, and this cyclical process might be supported by a variety of mapping tools and note-taking resources.

## References

1. Agarawala, A., Balakrishnan, R.: Keepin it real: pushing the desktop metaphor with physics, piles and the pen. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montreal, Canada (2006)
2. Barreau, D., Nardi, B.A.: Finding and reminding: le organization from the desktop. *SIGCHI Bull.* 27(3), 3943 (1995), doi:10.1145/221296.221307
3. Chen, C., Carr, L.: Trailblazing the literature of hypertext: author co-citation analysis (1989-1998). Paper presented at the Proceedings of the tenth ACM Conference on Hypertext and Hypermedia: Returning to Our Diverse Roots: Returning to Our Diverse Roots, Darmstadt, Germany (1999)
4. Elmqvist, N., Tsigas, P.: CiteWiz: a tool for the visualization of scientific citation networks. *Information Visualization* 6(3), 215232 (2007), doi:10.1145/1375939.1375943
5. Farooq, U., Zhang, S., Carroll, J.M.: Sensemaking of scholarly literature through tagging. In: *CHI 2009 Sensemaking Workshop*, pp. 49 (April 2009)
6. Malone, T.W.: How do people organize their desks?: Implications for the design of office information systems. *ACM Transactions on Information Systems (TOIS)* 1(1), 99112 (1983)
7. Matejka, J., Grossman, T., Fitzmaurice, G.: Citeology: visualizing paper genealogy. Paper presented at the CHI 2012 Extended Abstracts on Human Factors in Computing Systems, Austin, Texas, USA (2012)
8. Modjeska, D., Tzerpos, V., Faloutsos, P., Faloutsos, M.: BIVTECI: a bibliographic visualization tool. Paper presented at the Proceedings of the 1996 Conference of the Centre for Advanced Studies on Collaborative Research, Toronto, Canada (1996)
9. Sellen, A.J., Harper, R.H.R.: *The Myth of the Paperless Office*. MIT Press, Cambridge (2003)
10. Zhang, X., Qu, Y., Giles, C.L., Song, P.: CiteSense: supporting sensemaking of research literature. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy (2008)