**Design considerations for looking up information from multiple sources**

## Summary

Information for data entry can be scattered across multiple sources, and people tend to interrupt their task to collect digital information as soon as they realise they need it, as it is presumed to be easy to access in the moment. Switching away from the data entry interface can be disruptive as it slows people down and can increase errors. So how can users be better supported in managing their information needs? Based on both findings from my studies and previous literature on information search and management, this chapter explores a range of design possibilities to better support people in managing interruptions to collect information.

## INTRODUCTION

The previous studies reported in this thesis have shown that when looking up information for data entry work, people adopt different strategies. Study 2 showed that people group paper sources and collect this beforehand, but digital information is looked up when needed. Affordances of digital sources are more hidden than affordances of physical artefacts (Sellen & Harper, 2003), and it is less visible where to get information from and how long it will take to find information. As a result, people interrupt their data entry task as soon as they need digital information. Study 4 and 5 showed that if people learn how long it will take them to access information, they adapt their planning strategies. If participants knew it took them a long time to access a source, they postponed looking it up and entered other information first. An issue highlighted in Study 2 is that people often do not know where to get digital information from, and do not know if it will take them a long time until they have already interrupted themselves. Furthermore, whereas in Study 4 and 5 people always needed to use the same information from the same sources, participants in Study 1 and 2 did not always know they needed information until they had started a task. As soon as a need emerged, they addressed this need immediately to not have to hold it in memory. Interruptions are disruptive for data entry in a number of ways. It takes time for people to resume the task, they may have forgotten where they were and enter information in the wrong fields (Brumby, Cox, Back, & Gould, 2013; Monk, Trafton, & Boehm-Davis, 2008).

A large body of work has looked at how people organise, manage and re-find information, in order to design appropriate information management tools (Dumais et al., 2003; Trullemans, Sanctorum, & Signer, 2016). However, most of these studies tended to focus on finding information as a main task, and not as a subpart of other work and how people schedule interruptions from work to look up information. Furthermore, the tasks studied to evaluate these tools usually involved organising documents that had to be used for a longer time. The type of data entry task central in this thesis is characterised by rapidly going in and out of many different types of information sources. The current chapter reviews prior relevant work on information management and search tools, and explores a set of design possibilities on how people can be supported in managing interruptions to look up information.

## EXISTING WORK

Switching between documents and applications to look up information is common for many activities beyond data entry work. For example, people need documents when writing a paper, or use emails, calendars and written documents to plan a project. Prior research has explored tools to support fragmented work has focused on information management, information search, and integration of information from multiple sources.

## Task and information management

To make it easier to re-find documents for a particular activity, some systems have looked at grouping windows and documents. For example, *TagFS* (source) allows users to tag documents and define groups. *GroupBar* (Smith et al., 2003) makes it possible to group windows needed for a task in the task bar. This can be particularly useful when resuming an interrupted task: the user can see which documents were used before leaving the task.

These tools offer the user flexibility in organising information sources in different ways, but come with a number of limitations. First, it assumes the user knows in advance what information is needed for which purpose. While some information needs are known in advance of the data entry task, it regularly occurs the user needs unexpected information. Second, categorising documents can be time-consuming, and people are often not willing to invest time doing so (source). Especially for data entry tasks where documents are only briefly needed, people may not make the effort to group information. Lastly, studies have shown that when people do make the effort to organise documents into groups, they often use inconsistent labels, making it difficult to re-find information later (Jones, Jiranida Phuwanartnurak, Gill, & Bruce, 2005).

## Information search

In addition to information management, other studies have focused on supporting information search. An issue with looking for information is that information can be scattered across applications, and users have to go in and out of these separately to search and find what they are looking for. To support re-finding information across different applications, Dumais et al. (2003) presented a tool called *Stuff I’ve Seen*. Users were presented with a unified search interface which they could use to search through information they had already seen before, such as emails and web pages. A user study, where participants installed the tool on their computer and used it for two weeks, showed that users used the search tools of individual applications less frequently and used *Stuff I’ve Seen* instead. The focus of the tool was to improve search, rather than the scheduling or reducing of interruptions from work to search. The user still had to switch from their main task environment to a separate tool, and create a search query or use filters to view relevant search results.People do not always know what to search for, as demonstrated by both previous literature (source) and Study 2 of this thesis. Furthermore, for familiar documents, the preferred way of navigation is often browsing, rather than searching (Bergman, Beyth-Marom, Nachmias, Gradovitch & Whittaker, 2008).

Whereas *Stuff I’ve Seen* only supported searching for digital information, *PimVis* was developed by Trullemans, Sanctorum, & Signer (2016) to allow search across both paper and digital sources. A graphical user interface presented a visualisation of documents, grouped according to the context in which they are relevant. Bookcases and filing cabinets were augmented with LEDs, which would light up if users selected a document in PimVis that was contained in these physical locations. By opening a document in *PimVis*, the user could see documents related to this document. *PimVis* was evaluated using the task of finding documents for writing a paper. As *PimVis* was a standalone application, users had to switch away from their current application, such as their text editor, and open the document in *PimVis* to view its related documents. Participants reflected that *PimVis* would be useful for archived documents. For so-called ‘hot’ documents, which are needed for short-term tasks in the moment, they would value seeing related documents in the environment they are currently already working in, rather than having to go to a separate application. In the user study, the grouping of documents as well as augmentation of physical artefacts were set up by the researchers. The primary aim was to see whether participants could easily navigate through *PimVis*. They were given tasks to find specific documents, such as ‘You want to read the paper called X, which is related to the paper called Y’. In practice, the user would have to pre-define in which contexts documents were to be used and how they were related to other documents, which has the same drawbacks as categorising and labelling documents as discussed above.

## Interruptions and delayed intentions (in progress)

One reason why people are not always able to organise information efficiently is because they may not know they need information until they have started a task. In Study 2, participants disrupted their data entry work as soon as they realised they needed certain information. Prior work on self-interruptions found that office workers often start new tasks before completing old ones (Czerwinski, Horvitz, & Wilhite, 2004; Jin & Dabbish, 2009). If people are able to keep track of tasks they need to perform, it may help them in deferring these tasks until a more convenient moment, rather than addressing them as they realise they need to be done (Jin & Dabbish, 2009). For example, an interruption between subtasks is less disruptive than an interruption in the middle of a subtask (Bondarenko, 2010; Gould, 2014).

Gilbert (2015) looked at people’s off-loading behaviour in both an experimental and naturalistic setting. Participants had to remember to perform an action later, and had the option to offload this intention or to keep it in memory. In both settings, a majority of participants offloaded these intentions when they had the option, and this significantly improved their performance. Additionally, in Study 3 of this thesis, where participants had to remember which blocks to drag to which location, a selection of participants placed blocks nearby what they though the correct location was, to not have to remember its location, and as a reminder to place them there later.

These findings suggest that if people have to memorise which information to retrieve, they may benefit from options to offload these information needs, and are able to effectively defer information subtasks until a convenient moment in the main data entry task.

## Context of task: documents ‘at hand’

Bondarenko & Janssen (2005) compared how information workers store paper and digital sources. One user need they found was that documents should be embedded in task-related context information, as it helps to resume a task after an interruption. In addition, in Bi & Balakrishnan's (2009) study on large and multiple display use, office workers felt more focused on the task and immersed in their work when surrounded by task-relevant documents. A limitation of most tools discussed so far is that the task window and information window are separate, and users need to switch between these. Microsoft Office’s new feature *TAP[[1]](#footnote-1)*  instead is a built-in add-on, which allows users to place relevant documents in a task pane next to their working document. The aim of the feature is to keep focus on document creation, rather than looking up information. The feature is presented as a task pane within a document, such as a text document or email, and contains an overview of documents that may be relevant to the current document. The task pane initially shows files that are most frequently used. If the pane does not show the documents that the user is looking for, there is also a search option within the task pane.

The feature can reduce interruptions from the task interface. However, the TAP feature is application-specific: the user is only able to include other Office documents, and not information sources such as websites and databases. Furthermore, it is mainly focused on re-using content from archived documents, and assumes the user knows which documents to re-use. It may be less suitable for situations where people do not know where to get information from.

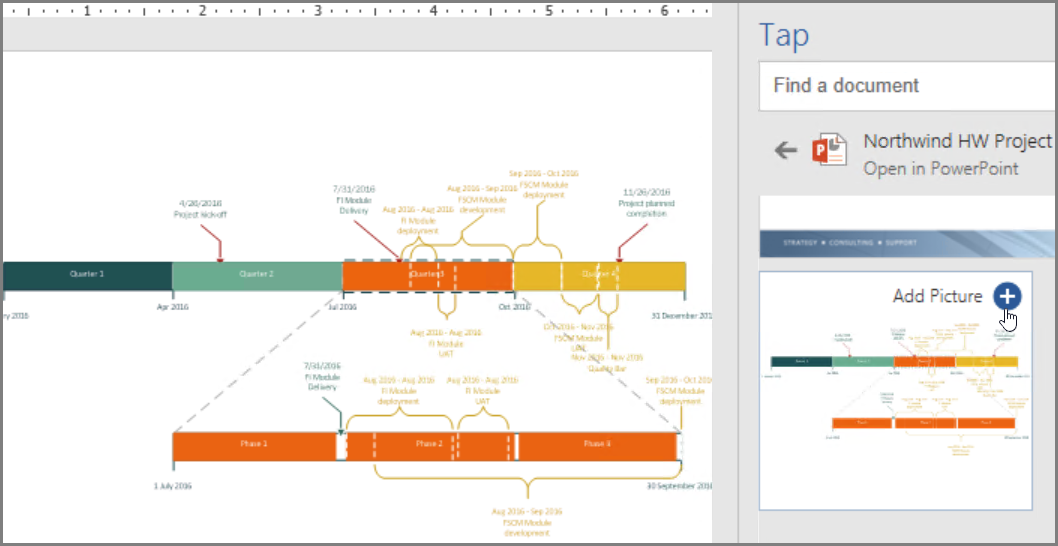


Figure 1. Microsoft's built-in feature TAP allows users to load and re-use content from other documents in a task pane.

## Type of activity

An important difference between previous work on information management and the studies in this thesis is the nature of the activity studied. Bondarenko & Janssen (2005) distinguish between two types of activities that information workers engage in: research activities and administrative activities. For an administrative activity, users go in and out of a variety of documents rapidly. For research activities, users need a smaller variety of documents, but these are needed for a long time. The tasks studied in most information management studies were more similar to research activities: participants had to read information to improve their understanding of a legal case (Cangiano & Hollan, 2009), or they needed to have the information for a longer time during a task (O’Hara, Taylor, Newman, & Sellen, 2002). A data entry task however is more similar to an administrative task. This distinction in activities is important, as it influences people’s information collection strategies (Bondarenko & Janssen, 2005)**,** and whether design considerations from previous work are applicable. Participants may not want to spend effort to organise information, such as grouping them and indexing them, if they only need the documents briefly. On the other hand, having contextual information nearby can be beneficial for both types of tasks, as it minimises interruptions and holding items in memory.

Table 1. Overview of previous solutions, and their advantages and drawbacks.

|  |  |  |
| --- | --- | --- |
|  | **Pros** | **Cons** |
| **GroupBar, TagFS** | All documents grouped in one place | * Need to categorise documents; suitable if use of documents is known in advance and doesn’t change; * even if known in advance, people may not know suitable names for categories or use inconsistent labels, and documents may be relevant in multiple categories |
| **Stuff I’ve Seen** | One unified interface to search across different applications | * Need to define search queries * people rather browse than search through familiar documents * Separate tool; user still needs to go out of task interface every time something is needed |
| **PimVis** | Search across both paper and digital sources | * Need to explicitly create groups of related documents, and and specify relations between documents * Separate tool; user still needs to go out of task interface every time something is needed |
| **TAP** | focus on document creation, not search, so there is no need to leave interface; ‘documents at hand’ | * only supports documents from the same application * documents have to be known in advance |

## Summary

Previous work on information search has looked at improving search across applications and media, but provides limited support for users on when to interrupt their work to conduct these searches. Prior work on information management has found that having contextual information at hand can reduce interruptions and helps users to be focused on their work. However, many of these tools require the user to organise, file and tag documents beforehand, and are based on the assumption that users know where to get information from. If people are able to off-load intentions to look up information, they may efficiently schedule when to interrupt their task and collect information.

## DESIGN CONSIDERATIONS

Based on both findings from the literature review and findings from my studies, the following requirements for information tools for data entry work can be defined:

1. Users should be able to group information for a task.
2. (Users should be able to search for different types of information sources.)
3. Users should be able to keep track of information they need.
4. Users should be able to off-load intentions for subtasks of looking up information.
5. The information interface should be embedded within the main task interface.

Three design alternatives are proposed below. For each design feature, Table 2 summarises which findings it builds on, and whether these findings are derived from previous research and/or my studies.

## Design 1: to-do list

**** ****

**** 

****

People may not always know they need information until they started a task. In this case, rather than interrupting it to look up information straight away, they can set a reminder for themselves to look this up later. People can write down information they need, to defer interruptions. It also allows people to offload memory and not hold information in memory. It does require people to write down more than they would normally and they still have to switch away from the data entry interface to look up information.

## Design 2: pinboard

Alternative two allows people to both write down information to look up, and pin the information sources in the information source. Once they have found information, they can pin this information source on their board, so if they need it again for future tasks, they can look it up and do not have to leave the interface.

## Design 3: search function



Alternative 3 allows users to search within the task pane for new information. While people do not have to leave the main task interface, it may be difficult to include all types of sources as people make use of web pages. Furthermore, people do not know what to search for.

Table . List of design features and the findings it builds on.

|  |  |  |  |
| --- | --- | --- | --- |
| **Design alternative no.** | **Feature** | **Builds on** | **Sources** |
| *1, 2* | *To do list:* by writing down information need for the task, people can keep track of, and off-load, intentions of information to look up | Research **suggested** people may be better able to defer interruptions if they are able to keep track of information they need  Research **showed** that when given the option, people offload and set reminders for delayed intentions  If people realise it takes time to complete a task, will delay a task but add visual reminder (e.g. a pile with unfinished claims)  People delay information needs they have to spend time on (looking up)  People have to retrieve information from both paper and digital; by having a to-do list they can schedule when to search for it | Jin and Dabbish (2009)  Study 3  Gilbert (2015)  Study 2  Study 4/5  Study 2 |
| *2, 3* | *Pin function.* If the user needs information repeatedly, there is option to pin information sources in the pane and have contextual task information nearby and ‘at hand’ when needed. | Previous research has **shown** that having context information nearby keeps people focused on the task  Previous work **suggested** that having context information nearby can aid resumption after interrupting the task.  If people need documents repeatedly, they make a local (paper) document to have a digital document ‘at hand’ | Bi and Balakrishnan (2009)  TAP  Bondarenko (2011)  Smith et al. (2003)  Study 2 |
| *1, 2* | *Flexibility of type of information sources.* The user has flexibility to look information up themselves in whatever way they want, whether through paper or digital sources. The tool does not aid them in searching for it, but rather aids them in scheduling when in the task to go search for it. | People do not always know what information they are going to need  People do not always know where to get information from  People use both paper and digital sources | Study 2 |
| *1, 2, 3* | *Offloading information.* When looking up information, the user can offload information to keep at hand when returning to the data entry interface. | People switch between various sources and hold items in memory | Study 1, 4, 5 |
| *1, 2, 3* | *Information pane.* The information pane to pin and write down information is embedded in the main task interface, so the user does not have to switch to a separate application to search or view information | … | … |

1. <https://support.office.com/en-gb/article/Find-and-use-the-content-you-need-when-you-need-without-leaving-Word-860118fc-1f61-41f6-922f-40084a284658> [↑](#footnote-ref-1)