Sewing the Seams of Sensemaking:

A Practical Interface for Tagging and Organizing Saved Search Results

Marti A. Hearst
University of California, Berkeley
102 South Hall, Berkeley, CA 94720
hearst@berkeley.edu

Duane Degler
Design for Context
4800 Hampden Lane, Bethesda, MD 20814
duane@designforcontext.com

ABSTRACT

This paper presents a usability-tested interface design that enables time-constrained analysts to organize their search results in a lightweight manner during and immediately following their search sessions. The research literature suggests that users want to lay out search results spatially in overlapping "piles," but a pilot study with a flexible canvas tool revealed that this design requires too much manipulation and has other drawbacks. This finding led to a novel hybrid design that combines structure with a flexible visual layout and which allows the analysts to quickly triage documents first and organize them later, or interweave these two processes. Two usability studies comparing the new design against a legacy tool found overwhelming preference for the new tool for saving and organizing search results. Design guidelines derived from this work could improve sensemaking interfaces for other search applications.

ACM Classification Keywords

H.3.3 Information Search and Retrieval: Miscellaneous

General Terms

Information Retrieval, User Interfaces

Author Keywords

Search User Interfaces, Sensemaking, Tagging, Organizing

INTRODUCTION

This paper describes the design and evaluation of a search user interface for a production setting, with real end-users who have complex, intensive, and time-constrained information analysis needs. This interface need resides at what we call a "seam of sensemaking," at the interstice between searching, saving, and organizing search results.

Ideally a sensemaking tool supports *flow* [7], allowing the user to fluidly and effortless move between operations such as querying, reading, saving, annotating, organizing, and labeling information [12, 28, 32]. However, most search interfaces either do not provide tools to both save and organize search results, or if they do, these tools go largely unused. The results reported in this paper provide design guidelines for how to build an effective user interface for both saving and organizing search results that was well received by target users (see Figure 1).

The task that this tool supports is a *triage* [16, 4, 3, 8, 21], allowing users to save documents into groups as they search, followed by (or interleaved with) an organization step in which the users work with the saved document groups and reorganize those groups and the documents they contain. These groups will be used for an analysis step downstream, and may also be re-used for later information needs. At any time during the search process, the analyst can pause in the search and annotate the groups with descriptive labels, remove or move documents, rearrange the groups, and so on, and then continue the search.

For these goals, the key design guidelines that can be derived from the results of the usability studies presented here are:

(1) Allow for immediate creation of and naming of new groups while documents are saved during the triage stage,

(2) By default, show groups in containers with an orderly spatial organization and a fixed size, to minimize the need to adjust sizes and positions of groups during triage,

(3) In the organization phase, provide grids or other guided support for moving objects, to minimize manipulation time, and

(4) Allow users to build searches based on saved groups.

The motivations and evidence for these guidelines will be explained in full below. Although much related work has been proposed to solve the problems addressed here, current designs and research prototypes do not support all four guidelines simultaneously.

This work is part of a larger project to design new search tools and interfaces for technical legal analysts who must determine which among millions of legal documents are the most relevant for their task, and use those documents to support a legal decision. The analysts generally have only a few hours for each decision-making task, using tools whose roots date back several decades but which nevertheless have powerful search features, in particular Boolean capabilities with proximity operators. Due to the time-constrained nature of their work, and past IT efforts that did not make use of user-centered design, the participants in the studies tend to be

HCIR '13, October 03 - 04 2013, Vancouver, BC, Canada

ACM 978-1-4503-2570-7/13/10$15.00.
http://dx.doi.org/10.1145/2528394.2528398

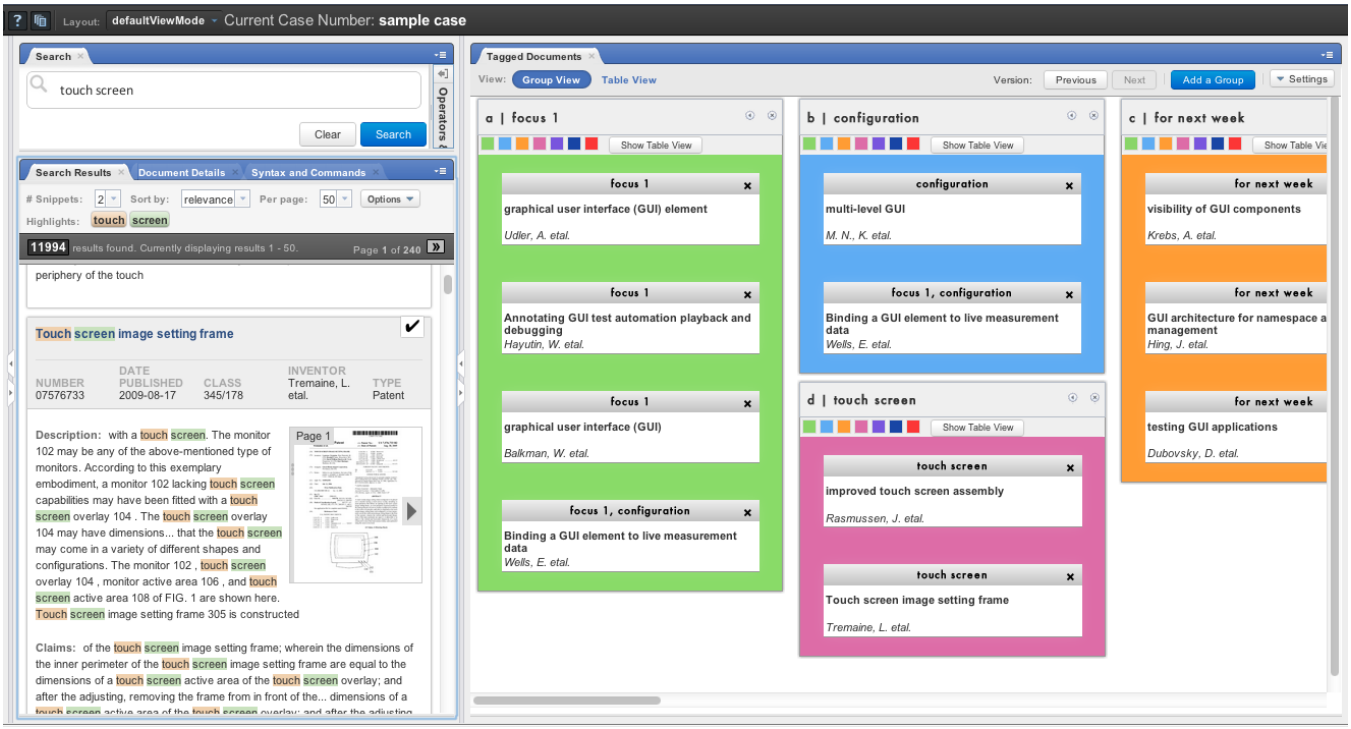


Figure 1. Design III: on the left are the query box and search results listing; the organizing interface is on the right. Four groups are shown with 11 selected documents; only the document titles are shown to save space. When a document is viewed in the search results, a keyboard letter (such as "a" automatically places it into a newly created group with the tag "a". For each group, the initial letter tag is visible (a through d) alongside the user-entered label added after the triage step. When documents are moved within a group, they "snap" into place in a grid and other documents are automatically moved as needed.

skeptical of new interface designs. However, the tools and interfaces for saving, organizing, and annotating saved documents in the legacy tools are not well-developed, and the users have expressed interest in seeing these redesigned and improved.

Currently, after narrowing the documents down to a smaller relevant subset, in most cases the analysts print out the documents, label and take notes on the paper versions, and organize them into piles or stack them on bookshelves for future use. The new software under development leverages documents as structured text with an underlying XML markup, supported by integrating sophisticated annotation tools, reducing the need to print and moving most of the process online.

In the remainder of the paper, we place this work in the context of related work, and then describe the three rounds of design and evaluation in more detail. We then summarize how these results can apply to the needs of searchers more broadly, and discuss future directions.

RELATED WORK

The Sensemaking Process

Taken as a whole, the sensemaking process has many iterative steps each of which can feed back on itself. Related work and existing software tools cover different subsets of this process, including discovery, retrieval, interpretation, management, and sharing of information [26, 32, 28]. The information foraging loop in turn should support many functions not standard in search tools today, including providing an overview of the collection, keeping track of what has been already viewed, suggesting what to look for next, showing temporal relations, and providing comparisons [27]. Much has been written about information organization strategies. For instance, Barreau and Nardi [6] found that short term organizational needs can differ from longer terms needs. They defined three information types: ephemeral, working, and archived. This paper addresses the equivalent of working information.

The term *triage* is used in several different ways; for instance, Buchanan and Loizides [8] use it to refer to finding relevant information within a particular document. We use it in the sense of Badi et al. [3], who define it as "the practice of quickly determining the usefulness and relevance of documents in a collection of documents."

More precisely, this work focuses on one of the *seams* between parts of the sensemaking process: the saving of search results during the search triage and the subsequent organizing of and thinking about the saved information, and the interleaving of the two. The concern for our users was being able to seamlessly move between saving, viewing, and organizing search results with minimal physical manipulation while at the same time supporting a spatial layout of those results.

Spatial Layouts with Search Results Triage

Considerable research finds that people like to spread information out in front of them on a table or a desk and organize information into physical piles, or use the physical layout of a virtual spreadsheet to organize information [24, 20, 18, 19]. Much sensemaking research has explored how to support this kind of interaction, both physically and virtually, and how it can also serve as a memory aid [29, 22, 18].

However, most of this work focuses on organizing documents *after* a subset of documents has been selected, that is, after the triage step, whereas our focus is on interweaving these two steps together. For example, the TopicShop system [1, 2] addressed two problems: how to determine which web site is of interest from that site's metadata and then, given a set of sites, how to group that set into categories effectively. Piles were allowed to overlap, but their manipulation and layout are not described in detail.

The VIKI and VKB research efforts [5, 22] discuss the advantages of sensemaking tools to support partially-formed ideas, to allow for re-thinking and re-organizing thoughts; they are a major motivation for this work. They recommended that sensemaking tools allow users to flexibly arrange, re-arrange, group, name, and re-name information. VKB supports a hierarchy of two-dimensional workspaces that allow orderly organization of representations of documents, and so present a promising model for our problem, but usability studies were done within small preselected groups of documents as opposed to tying the tool to search interfaces [5], thus leaving open the question of how they would work at the seam.

SketchTrieve was a research search project that explored the informal spatial layout metaphor, allowing users to place search results on a canvas view and extract subsets of the results and view them side by side [13], but this very early tool was not examined with usability studies. Other systems allow for grouping of small numbers of search results spatially into piles, but do not assess this aspect of their design [19, 14].

In the Stanford DLITE project [9], all aspects of a search were reified as graphical objects, and the results of a query were assembled as small icons lining the periphery of a circular region. Users could drag and drop retrieved documents into other containers but organization into categories or groups was not supported.

In the realm of multimedia search, VIGOR for video [11] and ImageGrouper for images [23] used drag-and-drop to form groups of objects that then become the query, which is especially helpful for items for which text querying is less rich. ImageGrouper emphasizes flexibility in the grouping, requiring a rectangle to be drawn around a relevant group, allowing overlap and requiring additional clicks to label a group.

Perhaps the most advanced tool for connecting search to a spatial layout, and the inspiration for the first round of design in the research reported here, is the combination of the TRIST search tool [16] and the Sandbox sensemaking tool [35], both designed for intelligence analysts. TRIST allows for thousands of search results to be viewed as icons in one display, supporting multiple linked dimensions that allow for comparing documents retrieved in response to different queries. It is also tightly coupled to the Sandbox, allowing the analyst to simply drag a document's surrogate into a flexible canvas workspace which includes the ability to create and name groups, to embed groups within groups, and to build templates that reflect hypotheses to be supported with evidence from the search results. To reduce time needed to make space in the canvas, a simple up-arrow gesture causes space to be cleared at the indicated location, and one item can knock away other objects; similar to QuickSpace [15] where an object expands, pushing others away. Sandbox has been proven successful with intelligence analysts working on complex synthesis problems. However, as will be discussed below, this kind of interface requires too much manipulation for the time-constrained workers under study here, and so perhaps is not as widely suitable for quick organizing as the design proposed in this paper.

On the other hand, some studies have cast doubt on the efficacy of using spatial layout with overlap in information organization interfaces. Rodden et al. [30] reported that participants in a study of layout of images disliked overlap and so modified the layout to be a uniform grid. In a recent study, Kodagoda et al. [17] found that users expressed dislike of overlapping cards, although many did like the freeform nature of the organizing tool. Fisher et al. [10] also noted the use of grid-like layout rather than anything more freeform when participants were asked to organize web clipping information.

Commercial Tools for Search Results Saving

Despite the interest in the research literature for freeform note-taking interfaces as part of the exploratory process, commercially available search tools do not easily support management of selection following the search itself. For instance, Google Notebook allowed users to save web pages and notes into collapsible folders directly from the search results and then reorganize them. However, a study by Russell et al. [31] found little to no sensemaking activity done with the tool. They note this may be due to design problems with the tool, so it is difficult to know if users do not want such a tool or if the execution caused it to fail.

Although Google Scholar[1] and Microsoft Academic Search do not provide search result saving functionality, academic online library catalog tools have traditionally supported saving searches and records of documents resulting from searches. Worldcat, a modern academic library catalog, provides a streamlined version of the most standard approach. A user must create an account and then create one or more collections in advance. After issuing a query, the search results list is shown with a checkbox alongside each search result. The user selects documents to save via the checkboxes, selects a collection from a drop-down box, and then clicks save. If the desired collection does not yet exist, the user is prompted to provide a new collection name and save to that. Items within collections can be assigned tags and exported

[1]Links for the sites mentioned in this paragraph are: **scholar.google.com**, **academic.research.microsoft.com**, **worldcat.org**, and **www.ncbi.nlm.nih.gov/pubmed**.

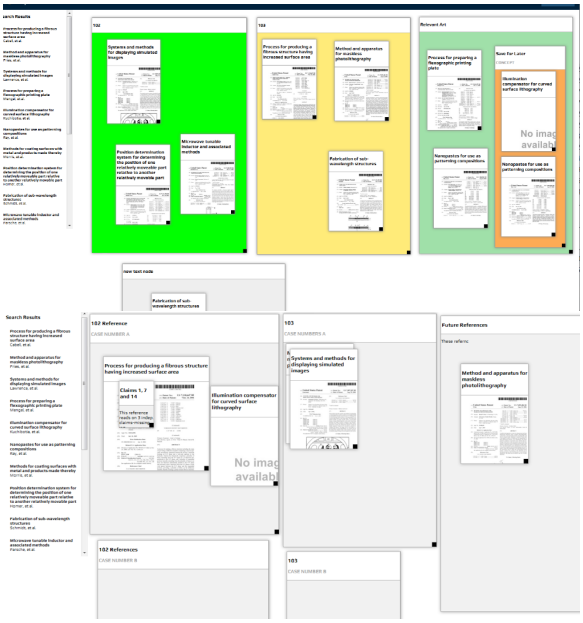


Figure 2. Two views of Design I.

to CSV files, and collections can be shared. Other citation-oriented search engines such as PubMed have similar facilities but their interface flows are more cumbersome and require far more operations to achieve the same functionality.

Also in the realm of academic citations are the Mendeley and Zotero citation storage and sharing tools. These are not search engines explicitly but rather are used in tandem with search engines and web browsers via web plugins and toolbars, allowing documents to be found anywhere and included into stored bibliographies. Both Zotero and Mendeley allow filtering by user-defined tags as well as by collections, but do not support more sophisticated sensemaking activities.

SUMMARY OF USABILITY STUDIES

The studies reported here began in June 2010 with informal interviews about all aspects of participants' work processes. The search tool under development supports sophisticated Boolean queries, search history, faceted classification, and other features not described here. For the particular subproblem of designing the interface for tagging, saving, and organizing search results, starting in July 2012, a series of focus groups and interviews were conducted to elicit user needs in more detail. An initial pilot interface was designed and in August 2012 (I) a Pilot assessment was completed, reported below. Based on those results, a more detailed interface prototype was developed and assessed in September 2012 in (II) The Prototype Assessment. The results of this study were quite positive; 17 out of 20 participants found it either "better" or "much better" that their current tool. In January 2013, after further refining the implementation, usage was evaluated in more depth by asking 10 new participants to use the tool on their own, outside of the lab in (III) the Self Paced Assessment. 9 out of 10 were positive about the design, thus confirming that these usually skeptical users saw value in the use of this new design. The studies are described in detail in the following sections.

INITIAL INTERVIEWS

Most of the analysts in the target group perform a triage step. During initial searches, a set of documents, ranging in size from as few as 10 documents to as many as 1000 (with an average around 30) are marked as potentially interesting, to be looked at in more depth in a second pass. This marking is often done in the main legacy tool using keyboard number keys, and is called tagging.[2]

Many of these analysts (6/20) use only one distinct tag to mark interesting documents, in part because the legacy tool makes it difficult to keep track of what the numbers stand for, and in part because it is difficult to view and manipulate the results once tagged. Other analysts use two, three, and even four tags. Often tags are used to reflect the strength of relevance (ascending or descending, although often only a single "relevant" tag). Some use them to assign categories of meaning, but because it is difficult to remember these, most wait until the second triage pass to make these more semantic markings. After the first triage tagging, some analysts will search within their tagged document set to narrow the quantity down, and re-tag in this second pass. Many (12/19) "untag" individual tagged documents after looking at them more closely. After this, most analysts (19/20) then print out the most relevant documents and write notes on them and highlight important passages. Some organize them semantically into groups. A few use a software spreadsheet to keep track of the combinations of categories for classifying each document.

Most analysts refer back to key documents that have been used in previous cases. Experienced analysts can have large collections of saved documents, but the legacy system does not support a way to integrate these with the rest of their work processes or with the search tools. Without exception, the analysts said that given a better tagging system, they would use it to keep track of potentially relevant documents.

In this first phase of design, analysts were shown alternatives. Several reacted positively to tools with a spatial layout like those seen in card sorting tools such as OptimalSort [33]. They reacted negatively to tools that allowed documents to overlap on one another; there was a desire for all documents to be visible. Differing views were expressed about whether the layout needed to be in grids or if a more free-form layout was feasible; analysts expressed a need to try it for themselves in order to really know which they would prefer. A few analysts expressed interest in nodes-and-links or network diagrams but most preferred a simpler layout with just boxes or groups. Throughout these and all interviews, the analysts emphasized the time-pressed nature of their work and their unwillingness to adopt new techniques unless they are truly useful, really save them time and effort, and enhance their work needs.

[2]The numbers reported in the this section are drawn from the 20 participants in study II.

PILOT DESIGN & ASSESSMENT (I)

Design

Based on the feedback from the focus sessions, we developed a rudimentary canvas tool that allowed the analysts to create groups consisting of rectangular boxes of any size, create groups within groups, place document surrogates within those boxes, resize, label, rename, and remove boxes and document surrogates. Figure 2 shows two examples.

To ascertain if this design was on the right track, a very small study, in the spirit of "discount usability testing" [25] was done with 7 analysts. If the design was well-received in its current form, this would be followed up with more participants. As it turned out, the design needed more work before further testing, as described next.

Procedure

Seven analysts were asked to revisit a recent case they had analyzed, and describe how they had tagged or otherwise kept track of the documents they found during their search. They were asked if they printed documents, how they organized those documents, and how they annotated or otherwise wrote on the paper versions of the documents.

Next the participants were asked to use the prototype organizational tool. First they viewed the tool in a starting state in which three blank boxes were shown. They were asked to manipulate the tool in various ways to learn its functionality. This included creating new boxes, resizing the boxes, placing boxes within boxes and then moving them around, and editing the text within the boxes.

Then the tool was reset to its original state with a view of three empty boxes, but this time a list of candidate documents was shown on the left hand side (see Figure 2), showing the results of a hypothetical search. Clicking on a document's title brought up an image of that document and placed it in the canvas; from there it could be moved and placed within a box.

Results

Four of the participants were strongly positively inclined towards the main ideas behind the tool, while 3 were ambivalent or felt they would not use such a tool. For all 4 of those who were positive about the tool, the spatial layout for organizing results made sense and seemed natural. Two immediately took to the tool, understood how it would be used, and had little or nothing negative to say about it. One of these participants requested a way to place links between the boxes to indicate relationships. The 2 others who were favorable made a number of suggestions for improvements. The positive aspects included:

- The "organic" feeling of the design
- The smooth interaction for resizing the boxes
- The way that inner boxes "stick" inside the enclosing box.

However, the analysts had numerous concerns, particularly about the amount of work needed to manipulate the view. They also requested features that would allow rapid switching between views of the retrieved documents.

In detail, these comments included:

- Too much work to resize, rearrange within boxes,
- Need a way to "snap" surrogates into line,
- Need to be able to see different views of documents with one click (figures, metadata, etc.),
- Need to be able to quickly switch between different overviews of groups (lists, boxes),
- Keyboard commands to assign labels would be useful, so documents would just appear in the groups, or in one group which the user later re-organizes,
- Should be able search within the contents of a box,
- Would like a view that automatically perfectly tiles and fits all the documents in a grid.

Additionally, participants made comments about bugs with the implementation (pressing "enter" within the text box did not always complete the edit change), or about designed features that were not yet implemented (could not drag directly from the search results into the boxes, had to click first, cannot view the full document, etc.).

Although initially attracted to the organic layout, the participants in the end wanted something neater that required minimal arranging on their part, at least until they had further reduced the number of documents sufficiently so that a more flexible organization would be viable.

NEW DESIGN GUIDELINES (II)

As a result of these studies, we modified the design to better meet the analysts' concerns. Both the analysts' comments and a careful reading of the literature [30, 17, 10] as discussed above suggest that although an organic layout with flexible, overlapping spatial layout has some appeal, that in fact to be practical for their needs the triage and organizational tools needed to exhibit the following properties:

- By default, an orderly organization of the containers for the spatial layout, a fixed container size to lend uniformity to the look, and control over the size of the boxes to reduce the need for scrolling and aid in navigation when gaining an overview of the groups.
- In order to save precious time and effort, minimal adjusting of the sizes and positions of the groups, at least initially during triage. Movement and reassignment are potentially acceptable at the organization and crystallization stage (which is also one point where deep annotation occurs), but should not be required at the initial triage and organization stage.

With these goals in hand, we designed a novel variation on the standard method for saving and organizing search results. This design supports a rapid triage followed by a more studied organization and analysis phase by default, but which nonetheless allows interspersing the two modes of triage and

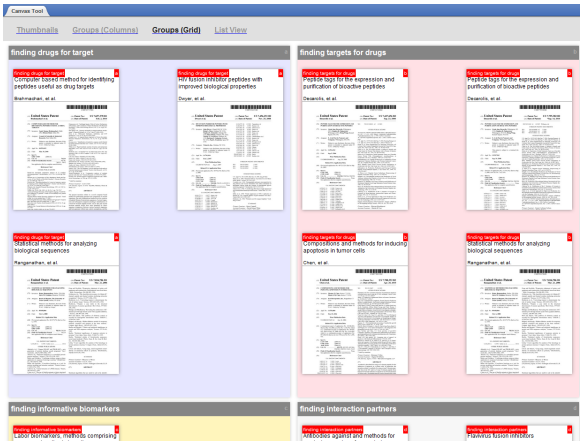


Figure 3. Fixed Grid View for Design II.

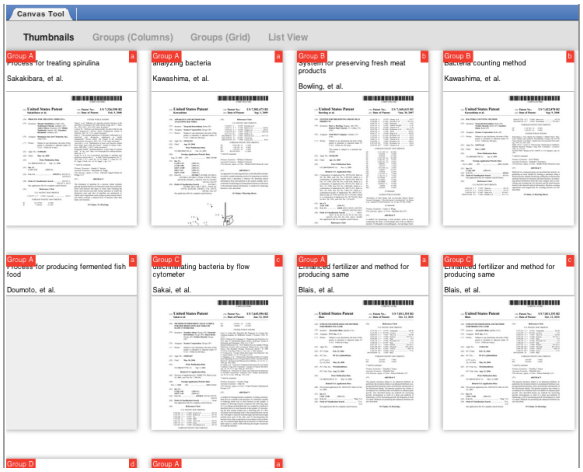
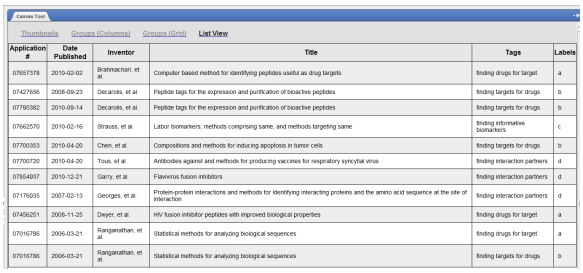


Figure 4. Thumbnail View for Design II.



Application #	Date Published	Inventor	Title	Tags	Labels
07657378	2010-02-02	Brahmachari, et al.	Computer based method for identifying peptides useful as drug targets	finding drugs for target	a
07427656	2008-09-23	Decarolis, et al.	Peptide tags for the expression and purification of bioactive peptides	finding targets for drugs	b
07795382	2010-09-14	Decarolis, et al.	Peptide tags for the expression and purification of bioactive peptides	finding targets for drugs	b
07662570	2010-02-16	Strauss, et al.	Labor biomarkers, methods comprising same, and methods targeting same	finding informative biomarkers	c
07700353	2010-04-20	Chen, et al.	Compositions and methods for inducing apoptosis in tumor cells	finding targets for drugs	b
07700720	2010-04-20	Tous, et al.	Antibodies against and methods for producing vaccines for respiratory syncytial virus	finding interaction partners	d
07854937	2010-12-21	Garry, et al.	Flavivirus fusion inhibitors	finding interaction partners	d
07176035	2007-02-13	Georges, et al.	Protein-protein interactions and methods for identifying interacting proteins and the amino acid sequence at the site of interaction	finding interaction partners	d
07456251	2008-11-25	Dwyer, et al.	HIV fusion inhibitor peptides with improved biological properties	finding drugs for target	a
07016786	2006-03-21	Ranganathan, et al.	Statistical methods for analyzing biological sequences	finding drugs for target	a
07016786	2006-03-21	Ranganathan, et al.	Statistical methods for analyzing biological sequences	finding targets for drugs	b

Figure 5. List View for Design II.

analysis, allowing the user who prefers to organize as they search to do so. Groups are again represented as boxes and arranged spatially on the canvas, but sizes are automatically adjusted to fit the document surrogates, and document surrogates are automatically placed and sized within the boxes. Boxes either grow to accommodate more document surrogates, or turn into a container with a scrollbar and a count indicating that they can hold more documents without increasing in the amount of screen size taken up. Boxes can still be nested. Each box has an editable label associated with it, and can be moved, resized, and deleted.

We designed the tool to meet additional properties that analyst requested in the pilot study:

- Allow for several views, and for seamless toggling among these views. These include at the very least a list view (Figure 5) well as spatial views.

- The spatial view allows the user to drag it to organize it, but also to allow for snapping into an orderly presentation, and for automatic space-filling actions.

- Three spatial views were contrasted: a Grid View in which 4 surrogates fit in each rectangle before scrolling starts (Figure 3), a Column View in which each surrogate appears in an ever-growing vertical group, and Thumbnail View (Figure 4) which does not group the surrogates.

The search portion of the tool was designed to allow for rapid, keyboard-based tagging to support the fast triage stage, as well as eventually drag and drop selection between the results and the canvas. Selection in results works synchronously with the spatial layout tool. As the document is tagged in the search results, it automatically appears in the canvas tool with the tag affixed. The list view (Figure 5) is sortable by metadata attributes such as date, author name, etc. It should also allow filtering by tag or label if they are assigned. Documents are automatically assigned spatial locations, depending on the view. Documents can be assigned into multiple tags/groups.

Some additional aspects of the interface include:

- If a new tag is used, a new group is created and is automatically assigned a location if none exists; the system is initialized with 4 empty groups to start,

- Groups can have labels assigned by editing their text label; the enclosed documents then have both the original tag and the user-assigned text label associated with them (a default label is assigned when documents are first tagged),

- Documents can be dragged and dropped into a group representing either a new or an existing tag, thus picking up the corresponding tags and labels,

- Within a group, the view of the document surrogates can be flipped easily (from thumbnail to list to keywords in context, etc.).

The studies were done on users' standard Windows machines with their dual-monitor (21", 1600x1200 resolution) side-by-side arrangement. Design decisions allowed the user to have maximum control over spatial layout for their own working style, creating ease of eye movement and attention.

PROTOTYPE ASSESSMENT (II)

The prototype design was evaluated by 20 target users in a laboratory-style study.

Study Design

Sessions lasted one hour. Participants were asked to bring a recent search record to the session, to give them ideas for searches to try out. First, the purpose of the study was

	Legacy tool	Prototype
Never	4	0
Rarely	5	4
Sometimes	8	5
Frequently/Often	2	10

Table 1. Changing Tags on Documents (out of 19 participants; note the shift to frequent changes with the new tool (Assessment II)).

explained, then they answered questions about their current search results tagging and saving practices. Next, the search tool and the four canvas views were demonstrated, and questions were answered. The participants then used the prototype to do their own searches for about 10-15 minutes and then answered questions on their opinions about the prototypes.

We anticipated that participants would prefer at least one of the canvas views over their current tool for tagging, due to the visibility of the resulting documents as they tag, and the ease of re-assigning tags. We also expected participants to anticipate using a greater variety of tags than with their current tool. Finally, we expected the analysts to prefer this design to their current tool. All three expectations were met.

Participants

20 analysts (8 female) participated in the study.

Preference for Grouping Strategy

After using the prototype, the participants were asked to rate the prototype in comparison to the legacy search tools. Using a scale of 1 (much worse) to 7 (much better), the overall rating was 6.1, and 85% (17/20) of the participants indicated grouping functionality was "better" (score of 6) or "much better" (score of 7).

Organizing Saved Results

Before the study began, participants were asked how often they change tags assigned to their documents using their current tool; results are shown in Table 1. After using the prototype, participants were asked how often they thought they would change tags using the new tool. The results show that the new tool supports a desire to change tags and is easy enough to use to enable this activity in future.

Participants commented that these sorts of activities are typically done on printouts of documents, rather than online, but they could see how this tool would change their work practices to be more electronic, and do more of the review and decision-making online.

Assessing the Usefulness of Different Views

Participants overwhelmingly felt they would use tags, given the capability that grouping offered them to use that information after selection. With the current tool, 6/20 assign only 1 tag per document, and one participant does not use tagging at all. As Table 2 shows, with the new design all participants anticipated using at least two tags. The numbers selected (2-5) was consistent with the way the design presented the groups of selections.

Preferences for Different Views

# Tags	Saving results	Organizing results
None	0	0
1	0	0
2-3	7	8
4-5	8	8
6 or more	5	3

Table 2. Estimates of how many tags would be used when saving (out of 20) and organizing (out of 19 participants) using the new tool (Assessment II).

	When saving	When organizing
Thumbnail	3	3
Groups (Columns)	12	8
Groups (Grid)	0	5
List	5	3
None of these	0	1

Table 3. Preferences between Canvas Layouts for task types (out of 20 participants) (Assessment II).

Participants were asked which views of the selected results would be most useful, both during the initial saving of interesting results and for organizing what they had saved. As Table 3 shows, in general, the Column view was preferred by participants, particularly for reviewing what they had tagged from the search results. Since 70% of participants (14/20) indicated they end up with 15 or fewer key documents at the end, this number will fit well within the column view. A task requiring larger numbers of saved documents may have revealed a different preference. Overall, the grouping of tagged results (in Columns or Grids) that is the basis of the design was preferred to the more traditional alternatives (60% during saving and 65% during organizing).

Comments about positive aspects of the designs include:

- The value of color coding groups for secondary visual cues, as well as labeling groups with attributes of their case or decision-making,
- The ability to move things between groups as their understanding of the information's relevance evolved over time,
- Controlling the order of items in a group, as a tangible reflection of their prioritization of information, which would be useful if they return to those references at a later time –whether a week or six months later.

The thumbnail view was not preferred by participants. In their comments, they mention:

- Positive: It provides a good overview of what they have tagged during their search session overall, and to find something without looking in groups (although they also suggest the list view could be better for this),
- Negative: It added little value over the grouped views, and did not confirm their tag selection as effectively. The view also did not allow them to apply their own labels, which help them think about the references they have collected for more detailed evaluation.

Keyboard Control vs. Drag and Drop

Only keyboard tagging was supported in the current prototype for moving documents from the search results to the canvas (which is what the analysts are accustomed to). Drag and drop did work for moving documents within the canvas tool, and participants were told that drag and drop was planned to be implemented for moving from the search results. When asked which mode they would use most often, 6/20 said keyboard commands, 4 said drag and drop, 9 said they would use both about equally, and 1 had no opinion.

Supporting Refresher Memory

Users often have to return to their saved reference material at a later date, as cases can move through multiple stages during their lifecycle, or when they encounter very similar cases that require similar research. Participants were asked about the usefulness of the group views and their close interaction with the information, to understand whether it might be useful to see the groupings when returning to that case research at a later date. On a scale of 1-7 (with 7 being "Very Useful"), the overall rating was 5.5 out of 7. For revisiting results, 75% of participants (15/20) rated the usefulness positively, with only 10% rating it negatively.

While this is something that can only be fully measured against actual use in their daily jobs, it is an important indicator of a potential benefit, given the challenges in time and productivity that they face now when using their current tools. Participants were asked overall, to the degree they felt the overall capabilities would be useful for their analysis tasks; on a 7-point scale the responses averaged to 6.1. (9 scores of 7, 7 scores of 6, 2 scores of 5, one neutral and one score of 3).

Desired Features for the Future

Participants asked for a number of additional features to enhance the capabilities. Some of the features they mentioned include:

- Allow each group panel to show either a thumbnail view or a list view, rather than having an entire view for one or the other,
- Allow items to be either moved or copied between groups, depending on how they think about the case they are working on,
- Searching within a group,
- Creating relationships between documents within a group or between groups, using some form of labeled connecting line, as a reminder supporting their synthesis of information inside documents,
- Changing what is represented in the thumbnail format of a document to make it easier to recognize what is important about it, such as a figure within a document, or a snippet from the search result, or a note applied by the analyst about the document,
- Have multiple canvases (sets of groups) that they can use over time, because they need different organizing schemes for reference information at different stages of their case – from initial review of results through different steps in the decision process, as well as revisiting cases months later if they are reopened.

SELF-PACED ASSESSMENT (III)

The results of the Prototype Evaluation (II) made clear the preference for the Column View. We further developed this design and continued to build out and mature other aspects of the tool's functionality. A screenshot of this final design can be seen in Figure 1. (The tool was written in java and javascript, using Lucene as the backend tool. The search supported complex Boolean queries with proximity operators, multi-color keyword in context highlighting, a scalable thumbnail image of retrieved documents in the search result, and optional relevance ranking using the default Lucene settings.) To assess the design in a setting with greater ecological validity, (that is, to further validate that this design has the potential to work for real users in a real setting) we arranged for 10 participants to test out the tool at their own pace and on their own search needs.

The other aspects of the search system did not allow participants to fully complete their work; only a fraction of their collections were indexed and the parser had not been fully refined. Therefore, they were asked to spend only 1 hour of their (costly) time using the tool.[3]

In an initial briefing session, participants were given a 30 minute demonstration of the entire search system as well as instructions about completing the study and filling out the post-search questionnaire. The portion describing the tagging and organization tool lasted about 5 minutes. Participants were told to search for documents across as many cases as they wanted within their 1 hour individual time allocation. Two participants completed 2 search cases in the allotted time, one completed 3, and the rest did one search apiece. The saving behavior varied from saving just one group with 3 documents to saving as many as 15 groups with 23 documents in two rows.

The comments about the strengths and weaknesses of the tool and other results of this evaluation closely mirrored those of Evaluation II. After using the tool for at least one real search task of their own choosing, 9 out of 10 participants felt that the core design worked well for their needs, assigning either a 6 or a 7 in a comparison to their current tool. The average score for the tagging functionality compared to the current tool was 6 out of 7. For using the organizing tool, the average score was also 6 out of 7.

Table 4 shows that 8 out of 10 participants felt they would use both the tagging and organizational features for most or every case. The two participants who indicated they tag "rarely" or "sometimes" using the current tool expressed enthusiasm about this design and indicated they will tag "for most cases" with the new tool.

[3]Originally 12 participants were recruited, but one became ill and another could not find any relevant documents to tag, so both were dropped from the study.

	Tagging tool	Organizing tool
For every case	7	5
For most cases	3	3
For a few cases	0	1
Never	0	1

Table 4. How often would you use this tool (out of 10 participants) for Assessment (III).

The one participant who did not find the tool useful felt that the grouping function was not helpful for them. Rather, this person only wishes to write a short annotation on the front of each document without grouping the documents into sets.

CONCLUSIONS AND FUTURE WORK

This paper described the iterative design and evaluation of a user interface at the seam between searching and saving and organizing search results. Three rounds of prototyping and usability testing with a total of 37 analysts were presented. The target users are legal analysts who do all of their work on an hourly time clock and they know that the tool being designed, when complete, is the tool that they will be required to use for their daily work. Therefore, they are highly motivated to give brutally honest feedback in these studies.

A major emphasis of the design is to increase efficiency and reduce the physical manipulations needed for creating new groups of selected documents when searching. The user can create a new group simply by typing a single character while in the search results: as the first document is selected, it is flagged by an icon in the results list where the user is focused. A corresponding group is created for that character, where the selected document automatically appears, in the periphery. This is in contrast with standard search saving interfaces where collections usually have to be explicitly created or are hidden in a different page or are shown as a textual list. As the user triages interesting documents, the documents are progressively added to the spatially distinguished groups, giving the user a visual sense of how many documents are appearing in each group and the characteristics of those documents.

As in Bae et al. [4], the use of multiple monitors and their impact on effective triage was a primary concern in this design. The design of the triage interface minimizes the transition switching between the displays or the levels of document detail, and the arrangement of groups provides primary and secondary focus areas. The simple keyboard tagging action allows the users' attention to remain on the text display during triage. Because the application is designed to easily span the dual monitors, the full interface is not just the organization area but a combination of results, organizer, and document viewer, so the analyst maintains focus but has all relevant alternative views of information near at hand.

Since these participants do the same analytical work daily, their estimates of their possible effort using a tool are most likely highly reliable. Not only did participants rate the functionality highly, but they predicted they would change tags frequently and otherwise use the tool to reorganize their saved searches. This is a highly desirable outcome because productive sensemaking usually requires reformulating of information structures [28, 32].

One area that will benefit from further study is to design specifically for longitudinal use by analysts [34]. Another is functionality to support saving documents to be used across cases and shared amongst analysts. Collaboration is an increasingly important aspect of the work, and currently they do this by sharing search query phrases and individual documents. Enhancing collaboration through sharing of organizing schemes and concepts is a potentially valuable next step.

As is often the case with designs initially targeted to those with special constraints, the resulting design seems promising and practical for a wider range of potential users. The guidelines proposed in this work should be considered by designers of other tools to better support rich sensemaking processes. For example, it is still difficult to plan a vacation using web search engines; better results saving and organizing tools could help when multiple scenarios and organizations are required. And for academic scholars, the problem of organizing research literature while searching can still evolve and improve.

Acknowledgements: This work was conducted to explore new concepts under the umbrella of a larger project at the U.S. Patent and Trademark Office. We thank Jaime Teevan for helpful comments on an earlier version of this work, and the anonymous reviewers for their suggestions which improved the paper.

REFERENCES

1. Amento, B., Hill, W., Terveen, L., Hix, D., and Ju, P. An empirical evaluation of user interfaces for topic management of web sites. In *Proceedings of CHI'99*, ACM (1999), 552–559.

2. Amento, B., Terveen, L., Hill, W., and Hix, D. Topicshop: enhanced support for evaluating and organizing collections of web sites. In *Proceedings of UIST'00*, ACM (2000), 201–209.

3. Badi, R., Bae, S., Moore, J. M., Meintanis, K., Zacchi, A., Hsieh, H., Shipman, F., and Marshall, C. C. Recognizing user interest and document value from reading and organizing activities in document triage. In *Proceedings of the 11th international conference on Intelligent user interfaces*, ACM (2006), 218–225.

4. Bae, S., Badi, R., Meintanis, K., Moore, J. M., Zacchi, A., Hsieh, H., Marshall, C. C., and Shipman, F. M. Effects of display configurations on document triage. In *Human-Computer Interaction-INTERACT 2005*. Springer, 2005, 130–143.

5. Bae, S., Marshall, C. C., Meintanis, K., Zacchi, A., Hsieh, H., Moore, J. M., and Shipman, F. M. Patterns of reading and organizing information in document triage. *Proceedings of ASIST'06 43*, 1 (2006), 1–7.

6. Barreau, D., and Nardi, B. Finding and reminding: file organization from the desktop. *ACM SigChi Bulletin 27*, 3 (1995), 39–43.

7. Bederson, B. Interfaces for staying in the flow. *Ubiquity 2004*, September (2004), 1–1.
8. Buchanan, G., and Loizides, F. Investigating document triage on paper and electronic media. In *Research and Advanced Technology for Digital Libraries*. Springer, 2007, 416–427.
9. Cousins, S. B., Paepcke, A., Winograd, T., Bier, E. A., and Pier, K. The digital library integrated task environment (DLITE). In *Proceedings of the second ACM international conference on Digital libraries* (1997), 142–51.
10. Fisher, K., Counts, S., and Kittur, A. Distributed sensemaking: improving sensemaking by leveraging the efforts of previous users. In *Proceedings CHI 2012*, ACM (2012), 247–256.
11. Halvey, M., Vallet, D., Hannah, D., and Jose, J. M. Vigor: a grouping oriented interface for search and retrieval in video libraries. In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, ACM (2009), 87–96.
12. Hearst, M. *Search user interfaces*. Cambridge Univ Pr, 2009.
13. Hendry, D. G., and Harper, D. J. An informal information-seeking environment. *Journal of the American Society for Information Science 48*, 11 (1997), 1036–1048.
14. Hinckley, K., Zhao, S., Sarin, R., Baudisch, P., Cutrell, E., Shilman, M., and Tan, D. Inkseine: In situ search for active note taking. In *Proceedings of the CHI'07*, ACM (2007), 251–260.
15. Hutchings, D. R., and Stasko, J. QuickSpace: new operations for the desktop metaphor. In *CHI'02 extended abstracts on Human factors in computing systems* (2002), 802803.
16. Jonker, D., Wright, W., Schroh, D., Proulx, P., and Cort, B. Information triage with TRIST. In *2005 International Conference on Intelligence Analysis* (2005).
17. Kodagoda, N., Wong, B. W., Rooney, C., and Khan, N. Interactive visualization for low literacy users: from lessons learnt to design. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 1159–168.
18. 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 (1983), 99–12.
19. Mander, R., Salomon, G., and Wong, Y. A "pile" metaphor for supporting casual organization of information. In *Proceedings of CHI'92*, ACM (1992), 627–634.
20. Marshall, C. C., and Shipman III, F. M. Spatial hypertext: designing for change. *Communications of the ACM 38*, 8 (1995), 88–7.
21. Marshall, C. C., and Shipman III, F. M. Spatial hypertext and the practice of information triage. In *Proceedings of the eighth ACM conference on Hypertext*, ACM (1997), 124–133.
22. Marshall, C. C., Shipman III, F. M., and Coombs, J. H. VIKI: spatial hypertext supporting emergent structure. In *Proceedings of the 1994 ACM European conference on Hypermedia technology* (1994), 13–3.
23. Nakazato, M., Manola, L., and Huang, T. S. Imagegrouper: a group-oriented user interface for content-based image retrieval and digital image arrangement. *Journal of Visual Languages & Computing 14*, 4 (2003), 363–386.
24. Nardi, B. A. *A small matter of programming: perspectives on end user computing*. The MIT Press, 1993.
25. Nielsen, J., and Hackos, J. *Usability engineering*, vol. 125184069. Academic press San Diego, 1993.
26. Paepcke, A. Digital libraries: searching is not enough. *D-lib Magazine 2*, 5 (1996).
27. Patterson, E. S., Roth, E. M., and Woods, D. D. Predicting vulnerabilities in computer-supported inferential analysis under data overload. *Cognition, Technology & Work 3*, 4 (2001), 224–37.
28. Pirolli, P., and Card, S. Sensemaking processes of intelligence analysts and possible leverage points as identified through cognitive task analysis. In *Proceedings of the 2005 International Conference on Intelligence Analysis, McLean, Virginia*, vol. 6 (2005).
29. Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D., and Van Dantzich, M. Data mountain: using spatial memory for document management. In *Proceedings UIST'98* (1998), 153–62.
30. Rodden, K., Basalaj, W., Sinclair, D., and Wood, K. Evaluating a visualisation of image similarity as a tool for image browsing. In *Information Visualization, 1999.(Info Vis' 99) Proceedings. 1999 IEEE Symposium on* (1999), 36–3.
31. Russell, D. M., Jeffries, R., and Irani, L. Sensemaking for the rest of us. In *CHI 2008 Sensemaking workshop* (2008).
32. Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. K. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (1993), 269–76.
33. Spencer, D. *Card sorting: Designing usable categories*. Rosenfeld Media, 2009.
34. Urban, J., and Jose, J. M. Ego: a personalized multimedia management and retrieval tool. *International Journal of Intelligent Systems 21*, 7 (2006), 725–745.
35. Wright, W., Schroh, D., Proulx, P., Skaburskis, A., and Cort, B. The sandbox for analysis: concepts and methods. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (2006), 801–10.