A Study of Visually Linked Keywords to Support Exploratory Browsing in Academic Search

Orland Hoeber and Soumya Shukla
Department of Computer Science, University of Regina, Canada

Abstract
While the search interfaces used by common academic digital libraries provide easy access to a wealth of peer-reviewed literature, their interfaces provide little support for exploratory browsing. When faced with a complex search task (such as one that requires knowledge discovery), exploratory browsing is an important first step in an exploratory search process. To more effectively support exploratory browsing, we have designed and implemented a novel academic digital library search interface (KLink Search) with two new features: visually linked keywords and an interactive workspace. To study the potential value of these features, we have conducted a controlled laboratory study with 32 participants, comparing KLink Search to a baseline digital library search interface modelled after that used by IEEE Xplore. Based on subjective opinions, objective performance, and behavioural data, we show the value of adding lightweight visual and interactive features to academic digital library search interfaces to support exploratory browsing.

Keywords: exploratory search, keyword visualization, academic digital library search, controlled laboratory study

Introduction
Searchers employ a variety of search strategies while looking for information within digital libraries, depending on the task type (simple or complex) and the features of the search system (Belkin et al., 1995). The success of a search process depends on how well the system supports these search strategies and how well it enables users to understand the impact of what they have done (Belkin et al., 1995). The study of the human element of search highlights the importance of good search interface design (Hearst, 2009; Wilson,

Corresponding author: orland.hoeber@uregina.ca

This is a preprint of an article published in the Journal of the Association for Information Science and Technology. This preprint has been updated to reflect changes in the final version. Please cite it as: Hoeber, O., Shukla, S. (2022). A study of visually linked keywords to support exploratory browsing in academic search. Journal of the Association for Information Science and Technology, 73(8), 1171–1191. https://doi.org/10.1002/asi.24623
A STUDY OF VISUALLY LINKED KEYWORDS

Providing interactivity beyond the traditional “query box and 10 blue links” interface pattern enables searchers to take an active role in the search process (Hoeber, 2008; Ruthven, 2008), which is critical when the information need requires the searcher to do more than just look-up facts.

While academic search tasks can span a wide degree of complexity (Hoeber et al., 2019), those that are classified as knowledge discovery may be considered complex search tasks, due to their open-ended and multi-faceted nature, and under-specification of the information need (White & Roth, 2009). Exploratory search has been proposed as a useful approach for tackling complex search tasks (Marchionini, 2006; White & Roth, 2009). Unfortunately, most publicly available academic digital library search interfaces provide little support for exploratory browsing, an important aspect of exploratory search. As a result, academic searchers are left to their own devices to perform exploratory browsing, and must develop their own strategies when undertaking the activities that have been highlighted as being critical to this aspect of exploratory search: discovery, learning, and investigation (White & Roth, 2009).

We hypothesize that if explicit support can be provided for exploratory browsing within a search interface, this may allow academic searchers to become better at finding relevant literature. This leads to two high-level research questions to be answered in this research: (a) how can we better support academic digital library searchers in exploratory browsing activities, and (b) what is the benefit of such approaches in comparison to a traditional academic digital library search interface?

To this end, we have designed and implemented a novel academic digital library search interface called KLink Search (Shukla & Hoeber, 2021) that uses visually linked keywords and an interactive workspace to support exploratory browsing. The visually linked keywords provide lightweight visualizations and simple interactivity to enable searchers to highlight keywords and use these to browse the search results. The interactive workspace uses this same visually linked keywords approach, enabling searchers to re-evaluate the relevance of the saved search results in relation to the highlighted keywords, as they refine their understanding of the search topic. In this paper, we explain how the features of KLink Search were designed to support exploratory browsing, and report on the design, conduct, and results from a controlled laboratory study that compared KLink Search to a Baseline interface (modelled after the IEEE Xplore interface).

Before we proceed, it is important to make the distinction between filtering and highlighting. Filtering techniques (e.g., faceted search features (Tunkelang, 2009) on metadata such as year, author, and publication venue) hide the search results that do not match the selected criteria, and primarily serve as a mechanism for building complex queries, rather than supporting exploratory browsing among the search results. Rather than filtering, KLink Search uses an approach that highlights keywords and their associated search results. The benefit of highlighting is that it supports comparisons and identification of gaps, both of which are important for the discovery, learning, and investigation aspects of exploratory browsing.

Related Work

This research was informed and motivated by three bodies of literature: exploratory search, exploratory search interfaces, and visualization-enhanced digital library search in-
Exploratory Search

When faced with a complex search task in which a searcher’s information needs are open-ended, multi-faceted, or ill-defined, search strategies that employ simple lookup approaches are not sufficient (Belkin et al., 2004). Instead, the exploratory search model has been proposed as a more useful approach (Jiang, 2014; Marchionini, 2006; White & Roth, 2009).

Building upon Marchionini’s early work on this topic (Marchionini, 2006), White and Roth suggested that exploratory search consists of two high-level phases: exploratory browsing and focused searching (White & Roth, 2009). Generally, a searcher undertaking complex search tasks will start with the exploratory browsing phase, and then proceed to focused searching when they have learned enough about the topic to reduce their uncertainty and focus on specific aspects of interest. This is akin to the transition between pre-focus and focus-formulation within Vakkari’s three-stage model of information seeking (Vakkari, 2003). There is also a possibility for the uncertainty to rebound as the searcher discovers new subtopics to explore (Jiang, 2014). See Figure 1 for an illustration of the exploratory search process and its relationship with uncertainty.

The goal of exploratory browsing is to develop an understanding of the information that is available for the current information need. The general pattern is for searchers to...
issue broad or general queries, and evaluate the search results to enhance their knowledge on the topic. Such a search strategy may be characterized as a breadth-first search; others have characterized this as heuristic-based (Jiang, 2014). Although the relevance of the encountered information might not be immediately apparent, as the searcher’s domain knowledge increases with the search process, their ability to identify relevant information improves.

White and Roth’s model specified three primary exploratory browsing activities: discovery, learning, and investigation (White & Roth, 2009). Discovery is a process of encountering and identifying new, previously unknown information. Learning is a process of internalizing and making sense of the discovered information, leading to the acquisition of new knowledge and understanding. Investigation is a higher-order cognitive activity that focuses on the analysis, synthesis, and critical assessment of what has been discovered and learned.

As a searcher becomes more certain about what it is they are seeking, they may transition into the focused searching phase. Here, the goal is to dive deeply into the specifics of the search topic. The general pattern is for the searcher to issue focused and specific queries, with a goal of finding precise information. Such a search strategy may be characterized as a depth-first search, and one in which analytic strategies are employed (Jiang, 2014). The ability to identify relevant information from irrelevant will not be difficult, given the focused nature of such searching.

White and Roth’s model specified three primary focused searching activities: query (re)formulation, result examination, and information extraction. Query (re)formulation is a process of specifying an initial query for the given focus of the search, and reformulating it as needed to isolate relevant information. Result examination is a process of carefully evaluating and assessing the relevance of individual search results for the current focus of the search. Information extraction refers to the activity of selectively identifying and saving relevant material from the information source.

A defining element during exploratory browsing, and the transition to focused searching, is the degree of uncertainty in the searcher’s understanding of their information need. When faced with a complex search task, a searcher will often have a high degree of uncertainty in what they are looking for and what they need to do. By undertaking exploratory browsing, and the associated activities of discovery, learning, and investigation, this uncertainty will be reduced, allowing the searcher to proceed to focused searching. It is also possible for the uncertainty to rebound as a searcher discovers new information (Jiang, 2014). Since the reduction in uncertainty is critical within exploratory search, we focus our work in this paper on methods to support exploratory browsing that enables discovery, learning, and investigation.

Exploratory Search Interfaces

It has long been established that while the standard list-based search engine results pages (SERP) may work well for lookup search tasks, they do not provide sufficient support for complex search tasks (Hearst, 2011; Hoeber, 2008; White et al., 2005). Such interfaces require that the searcher evaluate the results sequentially and in-order to assess their relevance in relation to their information needs and provide little interaction support for
in-depth information seeking strategies such as exploratory search. To address this shortcoming, a variety of search interfaces have been proposed to support exploratory search, following a general pattern of providing supplemental information to searchers in a format that encourages interaction beyond simply scrolling and clicking search results. While some have organized their reviews of exploratory search interfaces based only on how the search results are organized (Jiang, 2014), we do so along three dimensions: (a) the source of the supplemental information; (b) how the supplemental information is presented to the searcher; and (c) the type of interactivity that is supported.

A common source of supplemental information are the keywords associated with each individual search result. The source of such keywords may be from the most frequent within the search results (Hoeber & Yang, 2009), named entity recognition (Ahn et al., 2010; Krestel et al., 2011; Lin et al., 2010), or explicit author-assigned keywords (Kangasrääsiö et al., 2015; Ruotsalo et al., 2018). A unique approach is to extract such keywords from documents bookmarked by other users, supporting collaborative exploratory search (di Sciascio et al., 2018). Taking a step beyond keywords, some research has used explicitly selected text (with a reference back to the search result) and searcher-supplied annotations as the supplemental information to support exploratory search (Choi et al., 2021; A. R. Ward & Capra, 2021). Considering the query as a critical piece of information, explicit relationships between query terms and the search results may be extracted (Hoeber & Yang, 2008), or the query may be issued to other domain-specific repositories to obtain supporting or contextual information (Bozzon et al., 2010). Conversely, some have studied the use of multiple reusable sub-queries (lenses), producing supplemental information in the form of relevance scores with respect to each aspect of the query (Chang et al., 2019). Others have explored the extraction of metadata such as time (Hoeber et al., 2016; Krestel et al., 2011) and location (Krestel et al., 2011).

To a large degree, the mechanisms used to represent the supplemental information depend on the nature of the information and the need to show relationships. A common approach is to use colour encoding to show the relationships between keywords and other elements of the search interface, such as their contribution to the ranking algorithm (di Sciascio et al., 2018) or the association between sub-queries or query terms and search results (Chang et al., 2019; Hoeber & Yang, 2008). Radial layouts have been used to spatialize the information, showing similarity between objects (proximity), as well as importance (closer to the centre) (Kangasrääsiö et al., 2015; Ruotsalo et al., 2018). Simple tabular or grid layouts have also be used (Bozzon et al., 2010; Choi et al., 2021; A. R. Ward & Capra, 2021), along with tag clouds (Ahn et al., 2010; Lin et al., 2010), histograms (Hoeber & Yang, 2009), space-filling representations of weighted information (Chang et al., 2019), timelines (Hoeber et al., 2016; Krestel et al., 2011), sparklines (Hoeber et al., 2016), and maps (Krestel et al., 2011).

How the interactivity built into the search interface supports the exploratory nature of the search is a critical element. Saving documents into a workspace is common (Ahn et al., 2010; di Sciascio et al., 2018; Lin et al., 2010; A. R. Ward & Capra, 2021). Highlighting or filtering documents that match specific interactive selections (Ahn et al., 2010; Chang et al., 2019; di Sciascio et al., 2018; Lin et al., 2010) can help a searcher to isolate subsets of the search results; producing these as new searches that can be compared to the original can help the searcher to experiment during the search process (Hoeber et al., 2016).
Giving users interactive control over the ranking algorithm can provide searchers with the ability to specify what aspects of their search they find important at a specific point in the process (di Sciascio et al., 2018; Hoeber & Yang, 2008, 2009). Enabling searchers to drag-and-drop objects into a structured information space can give them a great deal of control in defining the query (Kangasrääsiö et al., 2015; Ruotsalo et al., 2018) or organizing saved information (Choi et al., 2021; A. R. Ward & Capra, 2021). For highly interactive interfaces, providing a mechanism for the searcher to predict the results of their interaction before they commit to the change can be very valuable (Kangasrääsiö et al., 2015). Interactively integrating search results from other domain-specific repositories can be beneficial when the search task is truly multi-faceted (Bozzon et al., 2010). Interactively layering on new sub-queries can help searchers to isolate different aspects of the search results they want to investigate (Chang et al., 2019). Visual representations such as timelines and maps lend themselves to pan and zoom operations, which result in the filtering of associated search results (Hoeber et al., 2016; Krestel et al., 2011).

Visualization-Enhanced Digital Library Search Interfaces

The field of information visualization deals with the design, development, and study of methods for visually encoding abstract data (M. O. Ward et al., 2015). The ultimate goal is to enable users to perceive, interpret, and make sense of the presented information (M. O. Ward et al., 2015; Ware, 2012). An important case for using information visualization techniques is to convey relationships among the data; the cognitive processes associated with inferring relationships among what we see is informed by the Gestalt Laws (Koffka, 1935). Within the information visualization literature, these have been translated into a set of design principles that can be followed to enhance a user's ability to see intended relationships within the visual display (Ware, 2012). Further, while there are a wide variety of methods for visually encoding data element properties (e.g., position, size, light/dark), colour/hue plays a vital role when the intention is to show similarity between related items, or the differences between those that are unrelated (Bertin, 1983; Ware, 2012).

Such information visualization techniques have been studied within the context of search interfaces for decades, with much of the early work not proving to be particularly effective at improving search performance (Hearst, 2009). However, as the concept of interactive information retrieval has developed in recent years Belkin, 2015, there has been a renewed interest in applying information visualization techniques to support the interactive search process (Hoeber, 2018), and in particular, the exploration of academic literature (Fedele et al., 2017). While the previous section presented a body of literature on exploratory search interfaces, here we focus specifically on interfaces designed for digital library search which use visualization as a fundamental component.

A common approach in visualization-enhanced digital library search interfaces is to provide a visual overview of the search results, which can be done using document spatialization techniques (Heimerl et al., 2016), showing keyword, tag, or topic distributions (di Sciascio et al., 2017; Khazaee & Hoeber, 2017; Medlar et al., 2016; Medlar et al., 2021) and spatializations (Peltonen, Strahl, et al., 2017), providing a timeline of when the documents were published (Dörk et al., 2008; Jackson et al., 2016), showing where the search results exist within an external knowledge structure (Clarkson et al., 2009; Sarrafzadeh & Lank, 2017), or providing compact summaries in a dense grid display (Bernard et al., 2015). Some
approaches visualize document-level meta-data such as citation patterns (He et al., 2019; Khazaeei & Hoeber, 2017; Matejka et al., 2012), bibliographies (Dattolo & Corbatto, 2018), temporal topic models (Medlar et al., 2016), or other attributes such as year, author, venue, location, or keywords (Bernard et al., 2015; Dörk et al., 2008; Kang et al., 2006; Khazaeei & Hoeber, 2017; Spindler & Dachselt, 2009). Others promote transparency by illustrating which aspects of the query contribute to the rank score of each search result (di Sciascio et al., 2017).

While the aforementioned methods work in conjunction with the SERP to support discovery and learning, visualization techniques may also be employed in the context of saved documents to support investigation. Many techniques for information organization and information re-finding have been proposed in the context of performing exploratory web search (Capra et al., 2010). Recently, workspaces have been used to organize important information separate from the SERP, storing relevant keywords (Peltonen, Belorustceva, et al., 2017) and annotations (Nedumov et al., 2019), tracking previous interactions such as queries and relations between the browsed pages (Park et al., 2008), and enabling cross-session search (Jhaveri & Räihä, 2005; Li et al., 2020).

Of particular relevance to the work in this paper is the prior research on using keywords to provide overviews of the SERP (Ahn et al., 2010; Hoeber & Yang, 2008; Kangasrääsiö et al., 2015; Krestel et al., 2011; Lin et al., 2010), approaches for illustrating the relationships among the search results (Chang et al., 2019; Hoeber, 2018; Khazaeei & Hoeber, 2017), and providing simple overviews of search intent models (Kangasrääsiö et al., 2015; Ruotsalo et al., 2018). One of the criticisms of such approaches is that the keywords are disconnected from their source documents, making them difficult to use when assessing and comparing the search results themselves. The approach we present in the following section was designed to address this shortcoming, and is used not only in the SERP but also in the workspace.

KLink Search

Our goal in the design of KLink Search was to develop interactive features that could support academic digital library searchers in conducting exploratory browsing activities. We devised a visually linked keywords approach to support discovery and learning by making it easy for searchers to identify relationships between keywords of interest and individual search results. We developed an interactive workspace that uses the same visually linked keywords approach to help the searcher to investigate among the saved documents and continue to learn about their search topic. Considering the three dimensions we used to organize the literature on exploratory search interfaces, KLink Search uses author-supplied keywords as the primary source of supplemental information, and colour encoding to show relationships between keywords as well as between search results. Regarding interactivity to support exploratory search, KLink Search provides a workspace and the highlighting of search results via keyword selection. The details of these features are explained in the subsections that follow, along with discussions of how they can support critical aspects of exploratory browsing.
Visually Linked Keywords

Typical academic digital library search interfaces provide a standard “10 blue links” style search engine results page (SERP), which allows each search result to be considered one by one. The lack of interactivity, beyond the ability to view a selected search result, makes it difficult for a searcher to make use of what they have learned in evaluating prior search results. We have addressed this shortcoming by designing a method in which the keywords of previously selected search results can be used to help the searcher to readily identify both similar and different search results. Our general approach is to use information visualization techniques and the explicit selection of search results and keywords of interest, creating an approach we call *visually linked keywords*.

Visually linked keywords is a lightweight addition to the standard list-based SERP which shows relationships between search results based on their keywords (see Figure 2). These keywords are extracted from the metadata of each search result, and are provided in a list beside the typical paper details. When a searcher selects a search result they think might be relevant, all of the keywords for that document become clickable, which is conveyed to the searcher with an underline and the addition of a circle in front of each keyword. Those same keywords in all other search results are marked in the same way.

While this in itself can allow a searcher to scan the search results to find those that

![Figure 2](image-url)  
*Search results page of KLink Search, with one document selected and three keywords highlighted using three visually distinct colours.*
use the same keywords as the selected search results, a more detailed analysis is supported when particularly relevant keywords are identified and clicked. These are highlighted using a visually distinct colour within the circle markers, enabling the searcher to visual scan the keywords associated with the other search results to discover those that use the highlighted keywords. Such visual scanning can also reveal search results that do not use any of the highlighted keywords. If a searcher decides that a particular keyword is no longer of interest, another click removes the highlight.

When using perceptually distinct colours for the keyword highlighting, a relationship between the same keywords used in different search results will be perceived, as per the Gestalt Principle of Similarity (Koffka, 1935) (objects that visually resemble one another in terms of hue, shape, size, or other visually discernible attributes will be perceived to be related). If many keywords are selected in this manner, the ability to visually match the colours is diminished. There is also a limit on the number of perceptually distinct colours that can be shown in a compact graphical element (M. O. Ward et al., 2015; Ware, 2012). Given these visual constraints, we set the maximum number of highlighted keywords to nine, and visually encode these using a high-intensity colour map of perceptually distinct and unordered colours (see Figure 3).

Of note, the visually linked keywords remain active throughout the search session, enabling the highlighted keywords to support the evaluation of search results across multiple queries. This approach enables two key aspects of exploratory browsing: discovery and learning. If a particularly relevant document is found, it can be selected and relevant keywords can be highlighted. By doing so, similar documents can be easily discovered in the SERP using the aforementioned visual scanning approach. The dynamic nature of choosing which keywords to highlight, and then considering other search results that make use of these keywords, enables searchers to learn the relationships between search results.

Figure 3

The colour map and associated hexadecimal codes used for keyword highlighting, created using ColorBrewer (https://colorbrewer2.org/#type=qualitative&scheme=Set1&n=9).
Interactive Workspace

Workspaces have long been studied as a mechanism to support exploratory searchers to collect relevant artefacts (Furnas & Rauch, 1998; Shah & González-Ibáñez, 2010). While they are becoming increasingly common within academic digital library search interfaces (Jhaveri & Räihä, 2005; Li et al., 2020; Nedumov et al., 2019; Park et al., 2008; Peltonen, Belorustceva, et al., 2017; Peltonen, Strahl, et al., 2017), they generally operate as a container for saved search results or snippets, providing limited interactivity beyond adding and deleting items from the workspace, and in some cases supporting the sharing of work with others (Yue et al., 2014). While we could have done the same, we hypothesized that the visually linked keywords could also help searchers when they are re-evaluating the search results saved in the workspace.

To this end, we built a session-based interactive workspace that allows searchers to save documents across multiple queries as they undertake exploratory browsing. These are represented in the same format as the SERP, including the ability to continue to highlight keywords of interest, or unhighlight keywords that are no longer meaningful (see Figure 4). Such refinement of the visually linked keywords is also reflected in the SERP, enabling the searcher to leverage their current keyword interests in subsequent queries.

The documents in the workspace are presented in the order they were added to maintain the temporal consistency of the search process. This is important in helping the searcher to re-evaluate those documents that were found earlier in the search process, but may no longer be relevant as their understanding of their information need changes. Such documents can easily be removed from the workspace with a simple click.

Within the exploratory browsing process, the features in the interactive workspace support both learning and investigation. The same mechanisms within the visually linked keywords that enabled learning the relationships among the search results in the SERP also enable the searcher to learn the relationships among the search results saved in the workspace. Further, the ability to change the keyword highlighting within the interactive workspace supports the investigation of what has been found in the search session. In particular, a searcher may use these features to analyze whether the saved search results remain relevant to their current understanding of their information need, by only highlighting keywords that remain relevant. In doing so, clusters of documents may be identified and considered as a group, rather than individually.

Implementation Details

The underlying data used for KLink Search is pulled from the IEEE Xplore academic digital library. This digital library was chosen due to the availability of its API for accessing paper details, including the keywords. Two alternatives were available regarding the source of the keywords: those provided by the authors and those selected from a controlled vocabulary. KLink Search uses the author-supplied keywords, under the assumption that authors are in the best position to use keywords to describe the core contributions of their work. However, a case could also be made for the value of using the keywords selected from the controlled vocabulary. Such a change could be made with a trivial modification to the software. Further, changing the entire data source for KLink Search would also be possible, the difficulty of which would depend on the availability of gaining programmatic
The workspace in KLink Search (extended screenshot) allows searchers to continue to use the visually linked keywords to compare and contrast the saved search results. Here we can see that earlier search results focused on speech and speech analysis (blue and red), while later search results focused on magnetic resonance imaging (green), and many use support vector machines (purple).
access to the underlying search results data.

The KLink Search interface was implemented as a web-based application using HTML, CSS, JavaScript, and D3.js (Bostock, 2020). The back-end REST API uses Node.js (OpenJS Foundation, 2020a) and Express (OpenJS Foundation, 2020b). With each new query, the data is first scrapped from the IEEE Xplore interface (IEEE, 2019a), and then their API (IEEE, 2019b) is used to fetch the metadata for each search result. A caching system using MySQL (Oracle, 2020) ensures a consistent set of search results are provided while running the user study.

**Evaluation Methodology**

A variety of search interface evaluation methods have been proposed for the study of interactive information retrieval systems (Belkin et al., 2004; Borlund & Ingwersen, 1997; Kelly, 2009) and information visualization systems (Carpendale, 2008). The common goal is to study the extent to which the search interfaces assist users in their information seeking process. In this evaluation, we wish to make a direct comparison between KLink Search and another interface that includes features that are commonly found in existing academic digital library search interfaces. Given that the two core features in KLink Search are intertwined (the visually linked keywords are provided not only on the SERP but also in the workspace), a holistic study is required. The comparison is facilitated by a controlled laboratory study method.

In a controlled laboratory study, we manipulate the independent variables (search interfaces) to observe their impact on the dependant variables (subjective, objective, and behavioural measures). Variables not under direct investigation are normalized (e.g., task complexity, search result sets), allowing us to control the study environment and isolate the impact of the manipulation of the independent variables on the dependent variables (Kelly, 2009).

To support this study, we developed a search interface with the same core functionality of KLink Search but without the new features specified in the previous section, herein called Baseline (see Figure 5a). Both interfaces closely mimic the IEEE Xplore search interface, but without their additional query-focused features (e.g., collection filtering, temporal filtering, faceted navigation, query term highlighting) and other irrelevant interface elements (e.g., site navigation, advertisements), which would have served as confounding factors had we compared directly to that interface (see Figure 5c).

**Research Questions**

The following set of research questions were developed to guided the design of the study and assess the value of the exploratory browsing features implemented in KLink Search (visually linked keywords and interactive workspace). The implied comparison is to the Baseline, where these features do not exist, but everything else is the same.

**RQ1:** How is the perception of usability (ease-of-use, usefulness, and satisfaction) affected by the addition of exploratory browsing features?

**RQ2:** How is the perception of knowledge gain affected by the addition of exploratory browsing features?
Figure 5

The two search interfaces under investigation in this study (a and b), and the IEEE Xplore search interface (c) upon which the Baseline was modelled.

**RQ3:** How is the perception of uncertainty affected by the addition of exploratory browsing features?

**RQ4:** How is search performance influenced by the addition of exploratory browsing features?

**RQ5:** How is feature-level search behaviour affected by the addition of exploratory browsing features?
Study Design

Using a within-subjects design, each participant was exposed to both search interfaces. Two search tasks were provided to enable the participants to exercise the features of the search interfaces (more details on the tasks are provided in the following subsection). The exposure to the search interfaces and assigned tasks were varied using a Graeco-Latin square to counter learning effects. A third self-selected search task was used only with KLink Search, providing a method of verifying that the performance on the assigned tasks was consistent with that of one that is of personal interest and importance to the participants, and providing evidence that is more ecologically valid than we normally see in controlled laboratory studies.

To collect evidence to answer RQ1, the typical usability variables perceived ease-of-use, perceived usefulness, and perceived satisfaction were measured using a post-task questionnaire. Using the Technology Acceptance Model 2 (TAM2) (Venkatesh & Davis, 2000) as the basis, four questions were posed for each of these three variables, with data collected on a five-point Likert scale.

To answer RQ2, knowledge gain was operationalized as a combined effect in the change of a participant’s perceived knowledge and perceived interest. This was done to take into account that one’s interest in a search task and knowledge gain influence each other (Rotgans & Schmidt, 2017). Likewise, to answer RQ3, uncertainty was operationalized as a combined effect in the change of a participant’s perceived uncertainty and perceived confidence. This was done to take into account the complex relationship between confidence and uncertainty in information-centric tasks (Peterson & Pitz, 1988). Each of these four dependent variables were measured twice per task (pre-task and post-task questionnaires) using a five-point Likert scale.

To collect evidence for RQ4 (search performance) and RQ5 (search behaviour), data was collected via interaction logs and screen recordings. Time to task completion was calculated explicitly using task-based controls added unobtrusively in the search interface. The number of saved documents were extracted from the logs, and the specific documents were scored using a three-point relevance judgement assigned by both researchers independently. Specific feature use was also extracted from the logs to assess the search behaviour. The screen recording data was used to replay search sessions and validate the log data.

Simulated Work Tasks

Two similarly complex tasks were created to simulate the need for the knowledge discovery that is necessary in exploratory search (White & Roth, 2009), and were presented as simulated work tasks (Borlund, 2003). They were created following the advice provided by Kules & Capra (Kules & Capra, 2009): (1) tasks for exploratory search systems must indicate uncertainty that requires the discovery of new information; (2) the domain should be unfamiliar for the searchers, but they must find it interesting; and (3) the task description should provide enough context to the searchers enabling them to relate and apply the situation as required.

The specific tasks used in this study are provided in Figure 6. The similarity in task description was intentional, allowing us to maintain consistency in context and motivation between the two search activities. We verified that the complexities of the tasks themselves
Task 1: Suppose you are taking a graduate course on the topic of Virtual Reality. For this course, you need to write a research report on the virtual reality methods used in a specific domain. Your goal is to explore among the research literature to learn about the ways in which virtual reality is being used in specific domains, and then narrow things down to a single domain. You do not need to perform a deep or comprehensive search; instead your goal is to get to the point where you have a few articles that you can use as a start for a more focused search that you will perform later. Your starting query for this task is “virtual reality”.

Task 2: Suppose you are taking a graduate course on the topic of Wireless Networking. For this course, you need to write a research report on one of the research directions in wireless sensor networks. Your goal is to explore among the research literature to learn about the various different topics related to wireless sensor networks, and then narrow things down to a single topic. You do not need to perform a deep or comprehensive search; instead your goal is to get to the point where you have a few articles that you can use as a start for a more focused search that you will perform later. Your starting query for this task is “wireless sensor networks”.

Figure 6

Assigned academic digital library search tasks.

were similar, based on the similar number of search results returned after submitting the initial query.

Along with the task descriptions, instructions were provided to guide the participants through an exploratory browsing process. While asking the participants to follow such instructions may have caused some to search differently than they normally would, it had the benefit of guiding them away from using simple information seeking strategies when faced with the complex search tasks.

1. Issue the given starting query.

2. Evaluate the search results to find and save some papers that you think are relevant.

3. Review these search results you have saved, removing those that you no longer think are relevant.

4. Submit a new query based on what you have learned.

5. Repeat this process of evaluating and saving search results, reconsidering what you have saved so far, and re-querying to find new information.

6. You can repeat this loop as many times as it takes for you to feel confident that you have learned and saved enough papers to read them in detail later.

For the third search task, the same exploratory browsing instructions were provided. Each participant was explicitly instructed to conduct their search on a topic they are highly interested in, know little about, and have seldom performed searches on. As the task complexity, context, and motivation for searching could not be controlled for this self-selected task, it was not used for direct comparison with the Baseline.
Procedures

The study sessions were administered via video-conference (due to the COVID-19 pandemic). After collecting consent to participate in the study, a pre-study questionnaire was administered to collect demographic and prior academic search experience data. Training videos were played for each of the interfaces, illustrating their use in the context of performing an exploratory browsing task. Participants were given an opportunity to use each interface and ask questions about how they worked.

Once the participants self-identified that they were confident enough to start the actual tasks, we proceeded with the primary tasks and data collection phase of the study. Participants were given an assigned task and interface to use, and were asked to answer a pre-task questionnaire, conduct the search activity, and answer a post-task questionnaire. Three such task/interface pairs were provided to each participant, with the first two being assigned according to a Graeco-Latin square, and the last being the participants’ self-selected search task using KLink Search. Explicit stretch breaks were taken after the completion of each search task in order to avoid fatigue effects. The entire study process took approximately one hour. Participants were compensated with an opportunity to win one of 16 e-gift cards.

Participants

Participants were recruited from among our Computer Science graduate student population, as the search tasks were designed to target this population. We restricted this to current Master’s level students, those who completed their Master’s degree within the last year, and first year doctoral students. This user group represents searchers who understand the need to search within an academic digital library rather than searching the web for resolving academic information needs, but have limited experience using academic digital libraries (Hoeber et al., 2019).

Among the 32 participants, 40% identified as female; the rest male. 56% of participants reported that they conduct academic searches multiple times per week; all did so at least once per week. Regarding self-reported expertise, the participants were almost evenly divided among two groups: being extremely or very good at academic searching; or being moderately or slightly good. This distribution represents an accurate sample of the target population of graduate students in our program.

Data Analysis

For each of the concepts measured via questionnaires, multiple statements were provided and data was collected via Likert scales (four questions for each of perceived ease-of-use, perceived usefulness, and perceived satisfaction; two questions for each of perceived knowledge gain and perceived uncertainty). Although such raw data is ordinal in nature, aggregating the data for each concept allows statistical analysis to be done using ANOVA (Lantz, 2013), which is known to be a robust statistical analysis technique.

The search performance data (time to task completion, number of saved documents, relevance) are numerical in nature, and therefore can also be analyzed using ANOVA. The search behaviour data represented counts (occurrences of feature use by type), and therefore is numerical data. As such ANOVA was used to identify statistically significant differences
between how the interfaces were used. In all cases, the statistical significance level was set using a p-value of 0.05.

Results

The results of the study are organized by the five research questions posed in the previous section, all of which seek to provide insight regarding the benefit of adding visually linked keywords and the associated workspace to a traditional academic digital library search interface.

RQ1: Perception of Usability

A series of statements were posed in the post-task questionnaire, focusing on three aspects of usability: ease-of-use, usefulness, and satisfaction. Overall, participants in this study found KLink Search easier to use, more useful, and more satisfying than the Baseline after performing the two assigned tasks. This finding was confirmed with the results on the self-selected task for which only KLink Search was used (see Figure 7).

An analysis of variance (ANOVA) of the data collected after each of the assigned tasks confirmed the statistical significance of the difference in each measure (ease-of-use: $F(1,62)=53.4$, $p < 1 \times 10^{-9}$; usefulness: $F(1,62)=91.9$, $p < 1 \times 10^{-13}$; satisfaction: $F(1,62)=55.7$, $p < 1 \times 10^{-9}$). These results show that the addition of the novel exploratory browsing features embodied in KLink Search had a positive effect on the perceptions of usability of the search interface.

RQ2: Perception of Knowledge Gain

Since knowledge and interest influence each other (Rotgans & Schmidt, 2017), we operationalized the concept of perceived knowledge gain as an increase in both of these measures. Before and after each task was performed, participants were asked to specify their degree of knowledge and interest in the assigned topic, collected via the pre-task and post-task questionnaires, respectively. After using each interface on the assigned tasks, and then using KLink Search on the self-selected task, participants reported an overall increase in knowledge and interest, as shown in Tables 1-3.

Combining the measures and conducting ANOVA tests revealed statistically significant perceived knowledge gain after using KLink Search on both assigned tasks and the

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before</td>
<td>after</td>
<td>$\Delta$</td>
<td>ANOVA</td>
</tr>
<tr>
<td>Task 1</td>
<td>knowledge</td>
<td>2.56 (0.96)</td>
<td>3.31 (1.01)</td>
<td>+0.75</td>
</tr>
<tr>
<td></td>
<td>interest</td>
<td>3.75 (1.13)</td>
<td>3.94 (1.00)</td>
<td>+0.19</td>
</tr>
<tr>
<td>Task 2</td>
<td>knowledge</td>
<td>2.75 (0.93)</td>
<td>3.19 (0.75)</td>
<td>+0.44</td>
</tr>
<tr>
<td></td>
<td>interest</td>
<td>3.25 (0.86)</td>
<td>3.63 (0.72)</td>
<td>+0.38</td>
</tr>
</tbody>
</table>

Table 1

Mean values (standard deviation) for knowledge and interest when using the Baseline on Tasks 1 and 2 (larger values represent greater knowledge/interest; 3 is neutral).
self-selected task, but not after using Baseline. The degree to which knowledge was gained for Task 2 was greater than Task 1, which we attribute to subtle differences in the abilities of the participants to perform the tasks and the quality of the search results provided. These results confirm that the inclusion of the features to support exploratory search gave the participants an increase in their perception of knowledge gain, and that while knowledge gain can also be achieved when using the Baseline, it may be affected by other factors such as the skill of the searcher.

Of particular note are the results from the self-selected task (Table 3). While the

(a) Overall perceived ease-of-use.

(b) Overall perceived usefulness.

(c) Overall perceived satisfaction.

Figure 7

Comparison of the usability measures between the interfaces on the assigned tasks, and for the self-selected task while using KLink Search.
increase in interest was rather small, there was little room to improved from the interest collected before the search task was undertaken (mean value of 4.5 on a five-point Likert scale). On the other hand, the increase in knowledge was rather large (a full step in the five-point Likert scale). These findings are significant indicators of the value and benefit of providing the type of explicit exploratory browsing support in academic digital libraries embodied in KLink Search.

**RQ3: Perception of Uncertainty**

As with the perception of knowledge gain, we operationalized the perception of uncertainty using two concepts: uncertainty and confidence. Using the pre- and post-task questionnaires, participants were asked to specify their degree of uncertainty in the topic and confidence in their ability to search for information on this topic.

There was little change in these measures after using the Baseline, including a decrease in confidence in Task 1 after performing the search. However, after using KLink Search on the assigned tasks and the self-selected task, there was a self-reported decrease in uncertainty and increase in confidence (see Tables 4-6).

As these two concepts are opposite in nature, we invert the confidence in order to conduct the statistical analysis. Combining the measures and conducting ANOVA tests revealed statistically significant differences after using KLink Search on both assigned tasks and the self-selected task, but not after using Baseline. These results confirm that the inclusion of the features to support exploratory search allowed the participants to reduce their perception of uncertainty after completing the search tasks.

![KLink Search](image)

**Table 2**

<table>
<thead>
<tr>
<th>Task</th>
<th>Knowledge before</th>
<th>Knowledge after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>2.69 (0.70)</td>
<td>3.44 (0.81)</td>
<td>+0.75</td>
<td>F(1,31)=6.30, p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>3.75 (0.93)</td>
<td>4.19 (0.66)</td>
<td>+0.44</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3**

<table>
<thead>
<tr>
<th>Task</th>
<th>Knowledge before</th>
<th>Knowledge after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 3</td>
<td>2.69 (1.06)</td>
<td>3.97 (0.74)</td>
<td>+1.28</td>
<td>F(1,62)=20.2, p &lt; 1 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>4.50 (0.62)</td>
<td>4.69 (0.54)</td>
<td>+0.19</td>
<td></td>
</tr>
</tbody>
</table>

*Mean values (standard deviation) for knowledge and interest when using KLink Search on Tasks 1 and 2 (larger values represent greater knowledge/interest; 3 is neutral).*

*Mean values (standard deviation) for knowledge and interest when using KLink Search on Task 3 (larger values represent greater knowledge/interest; 3 is neutral).*
A STUDY OF VISUALLY LINKED KEYWORDS

20

Baseline

<table>
<thead>
<tr>
<th>Task 1</th>
<th>uncertainty</th>
<th>confidence</th>
<th>before</th>
<th>after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.75 (1.06)</td>
<td>3.69 (0.87)</td>
<td>2.63 (1.09)</td>
<td>3.56 (1.09)</td>
<td>-0.12</td>
<td>F(1,31)=0.00, p = 1.00</td>
</tr>
</tbody>
</table>

ANOVA

<table>
<thead>
<tr>
<th>Task 2</th>
<th>uncertainty</th>
<th>confidence</th>
<th>before</th>
<th>after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.81 (1.33)</td>
<td>3.44 (0.63)</td>
<td>2.31 (1.08)</td>
<td>3.56 (1.03)</td>
<td>-0.50</td>
<td>F(1,31)=1.08, p = 0.31</td>
</tr>
</tbody>
</table>

Table 4

Mean values (standard deviation) for uncertainty and confidence when using the Baseline on Tasks 1 and 2 (larger values represent greater uncertainty/confidence; confidence inverted for statistical analysis).

KLink Search

<table>
<thead>
<tr>
<th>Task 1</th>
<th>uncertainty</th>
<th>confidence</th>
<th>before</th>
<th>after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.88 (0.81)</td>
<td>3.56 (0.63)</td>
<td>2.00 (0.82)</td>
<td>4.25 (0.68)</td>
<td>-0.88</td>
<td>F(1,31)=13.8, p &lt; 0.001</td>
</tr>
</tbody>
</table>

ANOVA

<table>
<thead>
<tr>
<th>Task 2</th>
<th>uncertainty</th>
<th>confidence</th>
<th>before</th>
<th>after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.63 (0.96)</td>
<td>3.69 (0.87)</td>
<td>1.88 (0.89)</td>
<td>4.19 (0.66)</td>
<td>-0.75</td>
<td>F(1,31)=6.65, p &lt; 0.05</td>
</tr>
</tbody>
</table>

Table 5

Mean values (standard deviation) for uncertainty and confidence when using KLink Search on Tasks 1 and 2 (larger values represent greater uncertainty/confidence; confidence inverted for statistical analysis).

RQ4: Search Performance

Search performance was characterized along three dimensions: time to task completion, total number of documents saved, and precision of the documents saved at the end of the task. The time to task completion and number of saved documents was calculated from the interface logs (with the caveat that network connectivity issues occurred in two data collection sessions, resulting in the need to calculate the time and count the saved documents based on the video recordings). In order to calculate the precision, the two authors of this paper independently assessed the documents each participant saved for relevance to the tasks on a three-point scale (relevant, partially relevant, not relevant). There was a very high level of agreement between the two assessors (κ = 0.976). All cases of disagreement were resolved. We took a pessimistic view of relevance, collapsing the three-point scale into binary relevance judgements, with partially relevant assessments treated as not-relevant (Scholer & Turpin, 2009).

The results of these three measures are summarized in Table 7. We found that participants took more time to complete their search tasks using KLink Search, but not at a statistically significant level. This is expected, given that this interface included new features that would have been unfamiliar to the participants. In addition, for an exploratory search activity, spending more time on the task indicates a higher degree of engagement with the activity (see the following results on search behaviour), which is a positive outcome. While fewer documents were saved using KLink Search (not statistically significant), there
Table 6

<table>
<thead>
<tr>
<th>Task 3</th>
<th>uncertainty confidence before</th>
<th>after</th>
<th>Δ</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.97 (1.06)</td>
<td>1.81 (0.78)</td>
<td>-1.16</td>
<td>F(1,62)=28.2, p &lt; 1 × 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>3.81 (0.78)</td>
<td>4.47 (0.62)</td>
<td>+0.66</td>
<td></td>
</tr>
</tbody>
</table>

Mean values (standard deviation) for uncertainty and confidence when using KLink Search on Task 3 (larger values represent greater uncertainty/confidence; confidence inverted for statistical analysis).

Table 7

<table>
<thead>
<tr>
<th>Task 1</th>
<th>time to task completion</th>
<th>Baseline</th>
<th>KLink Search</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>539s (218)</td>
<td>600s (252)</td>
<td>F(1,30) = 0.605, p = 0.44</td>
</tr>
<tr>
<td></td>
<td>number of saved documents</td>
<td>4.9 (2.0)</td>
<td>4.0 (1.6)</td>
<td>F(1,30) = 1.88, p = 0.18</td>
</tr>
<tr>
<td></td>
<td>precision of saved documents</td>
<td>68% (0.24)</td>
<td>93% (0.15)</td>
<td>F(1,30) = 12.6, p &lt; 0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task 2</th>
<th>time to task completion</th>
<th>Baseline</th>
<th>KLink Search</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>530s (218)</td>
<td>556s (229)</td>
<td>F(1,30) = 0.105, p = 0.75</td>
</tr>
<tr>
<td></td>
<td>number of saved documents</td>
<td>4.9 (1.7)</td>
<td>4.6 (1.9)</td>
<td>F(1,30) = 0.157, p = 0.70</td>
</tr>
<tr>
<td></td>
<td>precision of saved documents</td>
<td>74% (0.35)</td>
<td>93% (0.11)</td>
<td>F(1,30) = 4.27, p &lt; 0.05</td>
</tr>
</tbody>
</table>

Mean values (standard deviation) for the search performance metrics.

was a statistically significant improvement in the precision of the saved documents when using KLink Search compared to the Baseline. This provides evidence that the participants were able to find and save documents that were relevant and compare these to one another in the workspace, enabling them to stay focused on the task at hand. Furthermore, the smaller number of documents saved using KLink Search, coupled with the greater precision of these documents, highlights the value of the interactive workspace and the ability of the participants to prune irrelevant documents as they learned about the topic (as per step 3 in the exploratory browsing instructions provided to the participants).

RQ4: Search Behaviour

An analysis of the search behaviour revealed a similar level of interaction between the Baseline and KLink Search, with respect to the query, show abstract, and access paper/add to workspace interactions (see Table 8). Among these, no statistically significant differences were found. Of note, to save a paper, participants using the Baseline were only able to click on it to open the paper in a new tab. Those using KLink Search had two options: clicking on the paper to open it in a new tab, or explicitly adding it to the workspace. Very few participants chose to view the actual document while using KLink Search, and instead predominantly chose to add papers to the workspace.
<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>KLink Search</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>F(1, 31) = 0.66, p = 0.42</td>
</tr>
<tr>
<td>query</td>
<td>3.50 (1.15)</td>
<td>3.94 (1.81)</td>
<td></td>
</tr>
<tr>
<td>show abstract</td>
<td>2.94 (4.40)</td>
<td>5.50 (6.50)</td>
<td>F(1, 31) = 1.70, p = 0.20</td>
</tr>
<tr>
<td>highlight keyword</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unhighlight keyword</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>access paper / add to workspace</td>
<td>8.88 (3.46)</td>
<td>11.1 (4.11)</td>
<td>F(1, 31) = 2.80, p = 0.10</td>
</tr>
<tr>
<td>access workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>highlight keyword from workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unhighlight keyword from workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>remove paper from workspace</td>
<td>2.56 (0.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>query</td>
<td>3.94 (1.48)</td>
<td>3.81 (2.17)</td>
<td>F(1, 31) = 0.04, p = 0.85</td>
</tr>
<tr>
<td>show abstract</td>
<td>3.00 (3.69)</td>
<td>2.44 (3.44)</td>
<td>F(1, 31) = 0.20, p = 0.66</td>
</tr>
<tr>
<td>highlight keyword</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unhighlight keyword</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>access paper / add to workspace</td>
<td>9.38 (4.40)</td>
<td>10.2 (4.59)</td>
<td>F(1, 31) = 0.26, p = 0.61</td>
</tr>
<tr>
<td>access workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>highlight keyword from workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unhighlight keyword from workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>remove paper from workspace</td>
<td>2.19 (0.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>query</td>
<td>4.94 (3.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>show abstract</td>
<td>4.38 (4.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>highlight keyword</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unhighlight keyword</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>access paper / add to workspace</td>
<td>8.66 (3.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>access workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>highlight keyword from workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unhighlight keyword from workspace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>remove paper from workspace</td>
<td>2.13 (0.87)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 8

Average frequency of feature use (standard deviation), grouped by interface and task.
Considering the new features that were added to support exploratory browsing, the feature-level behavioural analysis revealed that these were used (in some cases, extensively) on both the assigned tasks and the self-selected task. In the course of each search task conducted using KLink Search, approximately six keywords were highlighted and one keyword was unhighlighted from the SERP. Participants visited the workspace approximately twice per task, and removed about two papers that they had previously added to the workspace. From the workspace, participants intermittently used the keyword highlighting/unhighlighting (less than one interaction per task).

Clearly, there was a high degree of engagement with the exploratory browsing features in KLink Search on both assigned tasks, and the behaviour on the self-selected task followed a similar pattern. The extensive use of the visually linked keywords from the SERP showed that participants found them to be a valued addition. While they were used less on the workspace, their presence provided passive information about the relationships between the search results, enhancing the participants’ abilities to remove irrelevant search results from the workspace, the results of which we can see in the significant improvement in the precision of the saved documents.

Of note, the addition of these features did not adversely affect or impact the use of the core search features that were also present in the Baseline. These behavioural results corroborate the search performance findings in this study, showing why the time to task completion was higher and the final number of saved documents was lower after using KLink Search. They also add credibility to the subjective findings, showing that the novel features were indeed used prior to the participants answering the post-task usability, knowledge gain, and uncertainty questions.

Discussion

Keywords are an underutilized part of the metadata provided by academic digital library search engines. Our novel approach of using visually linked keywords, and making these also available within an interactive workspace, provides an example of how visualization methods can be used to enhance interactive information retrieval interfaces (Hoeber, 2018). The results from the user study provide strong evidence of the value of these features to enable exploratory browsing tasks (see Table 9 for a summary of the findings). Even though such features consume valuable screen real estate, there was clear evidence that they were used and contributed to positive subjective impressions and the ability to find highly relevant documents.

The findings from this study enhance our knowledge of exploratory browsing (as a necessary first step in exploratory search). They show that search interface features that are designed to support discovery, learning, and investigation will be used, and contribute to an overall decrease in uncertainty. This is important, as the decrease in uncertainty is what enables a searcher to transition from exploratory browsing to focused searching (White & Roth, 2009). To be able to do so with an inconsequential increase in time to task completion can be considered a success in terms of the lightweight visualization methods used to enhance the search interface.

The controlled laboratory study had the typical limitations that are due to the trade-offs between control and realism (high internal, construct, and conclusion validity, at the expense of external and ecological validity). Of note, controlling the assigned tasks
and exploratory browsing activities in the study allowed us to make direct comparisons between the interfaces assigned to the participants, at the risk of the participants not being personally invested in the outcomes of the search tasks. This limitation was mitigated with the use of a self-selected third task that was performed only with KLink Search, giving us data to verify the accuracy of findings from the assigned tasks, as well as insight into the potential value of the approach for ecologically valid tasks. Providing explicit instructions on how to undertake exploratory browsing served as a further limitation, as this may not be how the participants would normally undertake such searches. A future study in a naturalistic environment could reveal whether participants are able to make effective use of the features when given the flexibility to search how they please.

While the holistic nature of the study made it feel realistic, this was at the expense of making it difficult to isolate one specific feature as having had the most substantial con-

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Measure</th>
<th>Assigned Tasks</th>
<th>Verified with Self-Selected Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1: usability</td>
<td>ease-of-use</td>
<td>KLink Search *</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>usefulness</td>
<td>KLink Search *</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>satisfaction</td>
<td>KLink Search *</td>
<td>✓</td>
</tr>
<tr>
<td>RQ2: knowledge gain</td>
<td>increase in knowledge &amp;</td>
<td>KLink Search *</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>increase in interest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ3: uncertainty</td>
<td>decrease in uncertainty &amp;</td>
<td>KLink Search *</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>increase in confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ4: search performance</td>
<td>time</td>
<td>no difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>precision</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>saved documents</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>query</td>
<td>no difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>show abstract</td>
<td>no difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>highlight keyword</td>
<td>KLink Search</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>unhighlight keyword</td>
<td>KLink Search</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>access paper / add to workspace</td>
<td>no difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>access workspace</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ5: search behaviours</td>
<td>highlight keyword from</td>
<td>KLink Search</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>workspace</td>
<td>light use</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>unhighlight keyword from</td>
<td>light use</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>workspace</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>remove paper from workspace</td>
<td>KLink Search</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 9

*Summary of the key findings from the study that compared KLink Search to the Baseline. Note that * indicates statistically significant findings.*
tribution to participant success. In future work, an ablation experiment may be conducted to evaluate each novel feature in isolation, and against other potential approaches.

While neither the Baseline nor KLink Search included the faceted querying features that are becoming increasingly common in academic digital library search interfaces, this was done intentionally so that the study could isolate the value of the specific additions to the search results. As such features have been shown to improve both subjective and objective search performance (Tunkelang, 2009), combining such features with the visually linked keywords and interactive workspace developed in this research could have a significant impact on searcher engagement, satisfaction, and performance. An open question is whether the sum of the parts becomes less than the whole. That is, by adding too many interactive features (that have each been shown independently to add value), there is a risk that the interface will become overwhelming and visually unappealing, reducing its overall value. In future work, we plan to investigate these issues through further software development, deployment in a readily-accessible academic digital library search interface, and further user studies.

While we have shown the value of a particular approach for leveraging specific search metadata information which uses a specific visualization method, it is not the only metadata information that could be of interest to the searcher. Other metadata could also have been selected and presented instead of the keywords. Two avenues for future work are to investigate the differences between the precision of author-supplied keywords and the consistency of keywords chosen from a constrained vocabulary in relation to their use in KLink Search, and whether any other commonly available academic digital library metadata can be used to highlight the similarities and differences among search results in a similar lightweight visual manner.

Conclusion

Supporting search within academic digital libraries is an important research area, as it has the potential to improve the ability for researchers to learn about and find relevant literature for a given topic. While experts may have developed skills and strategies to be effective at such searching, less experienced researchers such as graduate students are often left to figure out how to find relevant literature on their own. The development of academic digital library search interfaces that provide explicit support for exploratory browsing, such as KLink Search, may help such searchers to become more effective at discovering new information, learning about their topic of interest, and investigating patterns and relationships among the academic literature.

The primary contributions of this work are two-fold: (1) it serves as an example of how lightweight keyword-based visualization features can be added to a standard SERP to support exploratory browsing, and (2) it provides subjective opinions, objective performance, and behavioural data that show the benefits such an approach can have on academic digital library searchers when they undertake exploratory browsing activities. We hope that this work will encourage other researchers to explore and study the benefits of creating visualization-enhanced academic digital library search interfaces that enable exploratory browsing activities, and will inspire the development of new academic digital library search interfaces that provide more interactive support for the knowledge discovery tasks they are meant to enable. We also encourage others to consider how similar approaches may be
developed to support the second stage of exploratory search (focused searching), keeping in mind that the activities that need to be supported are different.

While the design, development, and study of KLink Search was performed in the context of academic digital library search, the findings generalize to other settings in which complex search tasks such as those that require knowledge discovery are prevalent. Examples of such search settings are lawyers searching within patent collections, historians searching within news archives, and policy analysts searching within governmental document collections. A critical element in applying the findings from this research in such other search settings is that keywords or other metadata be available, which can be listed and visually linked to show similarities or differences among the documents.

Acknowledgements

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), through the Discovery Grant held by the first author (RGPIN-2017-06446).

References


A STUDY OF VISUALLY LINKED KEYWORDS

Information Science and Technology, 70(8), 843–857. https://doi.org/10.1002/asi.21171


A STUDY OF VISUALLY LINKED KEYWORDS


