This is a repository copy of *Exploratory search in digital humanities: a study of visual keyword/result linking*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/213010/

Version: Accepted Version

Proceedings Paper:
Hoeber, O., Harvey, M. orcid.org/0000-0001-5504-2089, Momeni, M. et al. (2 more authors) (2024) Exploratory search in digital humanities: a study of visual keyword/result linking. In: Proceedings of the Association for Information Science and Technology. 87th Annual Meeting of the Association for Information Science & Technology, 25-29 Oct 2024, Calgary, AB, Canada. Wiley-Blackwell , pp. 161-171.

https://doi.org/10.1002/pra2.1017

© 2024 The Authors. Except as otherwise noted, this author-accepted version of a paper published in Proceedings of the Association for Information Science and Technology is made available via the University of Sheffield Research Publications and Copyright Policy under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Exploratory Search in Digital Humanities: A Study of Visual Keyword/Result Linking

Hoeber, Orland University of Regina, Canada | orland.hoeber@uregina.ca

Harvey, Morgan University of Sheffield, UK | m.harvey@sheffield.ac.uk

Momeni, Milad University of Regina, Canada

Pirmoradi, Abbas University of Regina, Canada

Gleeson, David Northumbria University, Newcastle upon Tyne, UK

ABSTRACT

While searching within digital humanities collections is an important aspect of digital humanities research, the search features provided are usually more suited to lookup search than exploratory search. This limits the ability of digital humanities scholars to undertake complex search tasks. Drawing upon recent studies on supporting exploratory search in academic digital libraries, we implemented two visual keyword/result linking approaches for searching within the Europeana collection; one that keeps the keywords linked to the search results and another that aggregates the keywords over the search result set. Using a controlled laboratory study, we assessed these approaches in comparison to the existing Europeana search mechanisms. We found that both visual keyword/result linking approaches were improvements over the baseline, with some differences between the new approaches that were dependent on the stage of the exploratory search process. This work illustrates the value of providing advanced search functionality within digital humanities collections to support exploratory search processes, and the need for further design and study of digital humanities search tools that support complex search tasks.

KEYWORDS

Interactive information retrieval; exploratory search; digital humanities collections; user study; Europeana

INTRODUCTION

A distinguishing feature of digital humanities research is the reliance on digital records and computer technology to support the research process (Berry, 2012). A fundamental task in such research is to search within these digital collections to find records that support or refute a given hypothesis (Crymble, 2016; Gibbs & Owens, 2012). Such information seeking tasks may be complex in nature (multi-faceted, ambiguous, and open-ended), yet the tools provided follow a design pattern that is tuned for simple look-up search tasks. As a result, significant cognitive and administrative effort is expended when conducting such search tasks.

Exploratory search has been proposed as a process that is well suited to complex search tasks (Soufan et al., 2022; White & Roth, 2009). While the development of exploratory search interfaces for complex search in contexts such as academic digital library search has become popular (Gomes et al., 2022; Jackson et al., 2016; Medlar et al., 2016; Qvarfordt et al., 2014; Ward & Capra, 2021), it is unclear whether these approaches are transferrable to the unique context of searching within digital humanities collections, where the motivation for searching is different, the resources are more visual than textual, and the metadata associated with resources may not be as robust as in academic digital libraries.

The objective of this research is to study how a specific class of exploratory search interfaces (visual keyword/result linking (citation withheld to preserve anonymity during peer review)) can be used to search within digital humanities collections, and to assess their value for digital humanities scholars. This research was conducted in the context of an existing online digital humanities collection with a focus on cultural heritage: Europeana (Europeana, 2024a). The following two research questions guided the work.

> **RQ1:** How can visual keyword/result linking approaches be used to enhance the search interfaces provided by Europeana?
>
> **RQ2:** What is the value to searchers when adding interactive visual keyword/result linking approaches to the Europeana search interface?

The remainder of this paper is organized as follows. A literature review is provided to establish the foundational research that has informed this study. The approaches used for enhancing the Europeana search interface with visual keyword/result linking are presented. The user study methodology is explained. The results of the study are presented and discussed. Finally, the paper concludes with a summary of the contributions, limitations of the work, and an overview of future research.

LITERATURE REVIEW

As digital humanities scholars make extensive use of electronic resources (Sinn & Soares, 2014), use of search technology is fundamentally critical to their task (Crymble, 2016; Gibbs & Owens, 2012). Some have found that digital humanities scholars have a tendance to "cobble together ad-hoc solutions" using pre-existing tools (Trace & Karadkar, 2017). Recent work has highlighted the challenges that searchers face when they are trying to learn about a topic of interest, suggesting the need for more advanced search interfaces (Wu et al., 2024).

It is important to recognize that searching within digital humanities collections does not occur in isolation. Toms & O'Brien developed a model of technology needs of digital humanities scholars consisting of five phases: idea generation, preparation, elaboration, analysis and writing, and dissemination (Toms & O'Brien, 2008). While search may be used during the idea generation phase through "inspiration-oriented search" (Hill et al., 2016), it more commonly happens during the preparation phase, including both searching for resources and the synthesis of found material. This suggests the need not only for a typical search interface, but also a workspace in which to store and analyze what has been found.

Some have identified that as the search task pursued by digital humanities scholars becomes less precise, difficulties can arise (Buchanan et al., 2005). This move away from precision and towards conceptual searches suggests the move toward complex search tasks, which require searchers to employ strategies that go beyond simple look-up processes (Marchionini, 2006). Exploratory search has been proposed as a more useful approach in such situations (Soufan et al., 2022; White & Roth, 2009). Generally, searchers undertaking complex search tasks will start with an exploratory browsing style of searching, 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. When searchers are in the exploratory browsing style of searching, they are generally concerned with discovering, learning, and investigating aspects of their information needs. When transitioning to focused searching, the activities change to reformulating queries, examining results, and extracting information.

As the study of search interfaces for digital humanities collections has to date been little studied, we direct our attention to a related context: searching within academic digital libraries. Most current academic digital library search interfaces include features to support searching and saving documents to workspaces, relying on static lists to represent search results. Simply listing search results without providing a mechanism to enable searchers to make sense of them in relation to one another limits their ability to undertake exploratory search processes. Recent work on visual keyword/result linking provides an interactive mechanism that allows searchers to select keywords that are relevant to their current information need, enabling them to assess the relationships between items in the search results lists, both in the search results page and the workspace (citation withheld to preserve anonymity during peer review). We have extended this work to study multiple ways of visually linking keywords and results within a search results interface (citation withheld to preserve anonymity during peer review). These approaches serve as the foundation for the interfaces developed and studied in this paper.

ENHANCING THE EUROPEANA SEARCH INTERFACE WITH VISUAL KEYWORD/RESULT LINKING

Europeana is a digital multimedia repository that provides access to an extensive collection of cultural artifacts including art, historical documents, and multimedia resources (Europeana, 2024a). The search interface serves as a gateway to this collection of cultural artifacts for researchers, professionals, and the public (Hill et al., 2016).

The existing Europeana search interface follows a typical search interface design pattern for digital media, with a query box and a grid of search results showing an image, title, and source for each resource. Options exist to change the grid view to either a two-column list view (adding descriptions to each result) or a mosaic view (removing the title and source from each result). Facets are provided to allow searchers to filter the search results based on theme, type of media, and whether permission is required to use the resource. Mechanisms exist for "liking" resources or saving them into rudimentary workspaces, where they are presented in a four-column grid view only.

Baseline

A baseline search interface was constructed using the existing Europeana search interface as a template and the Europeana Search API as the data source (Europeana, 2024b). We instrumented the search interface, saving all user interactions with the LogUI library (Maxwell & Hauff, 2021). Small adjustments were made to the style of the interface, separating metadata information from actions for each search result (e.g., moving the save and like operations to the bottom-left region of the search result card) and removing features that might distract from the search focus of this study (e.g., collection suggestions). To maintain consistency within our study, a caching mechanism was implemented to ensure that each participant received an identical set of search results for the same queries. Figure 1 shows screenshots of the search interface and the workspace for the Baseline.

In order to answer RQ1, we considered the different ways in which visual keyword/result linking might be used to support search within the Europeana collection. We implemented two visual keyword/result linking approaches

which are described in detail in the sections that follow. These were built as extensions of the Baseline above, maintaining consistency between these interfaces as much as possible so that the only difference is the method for representing keywords and the approach for visually linking them to the results.

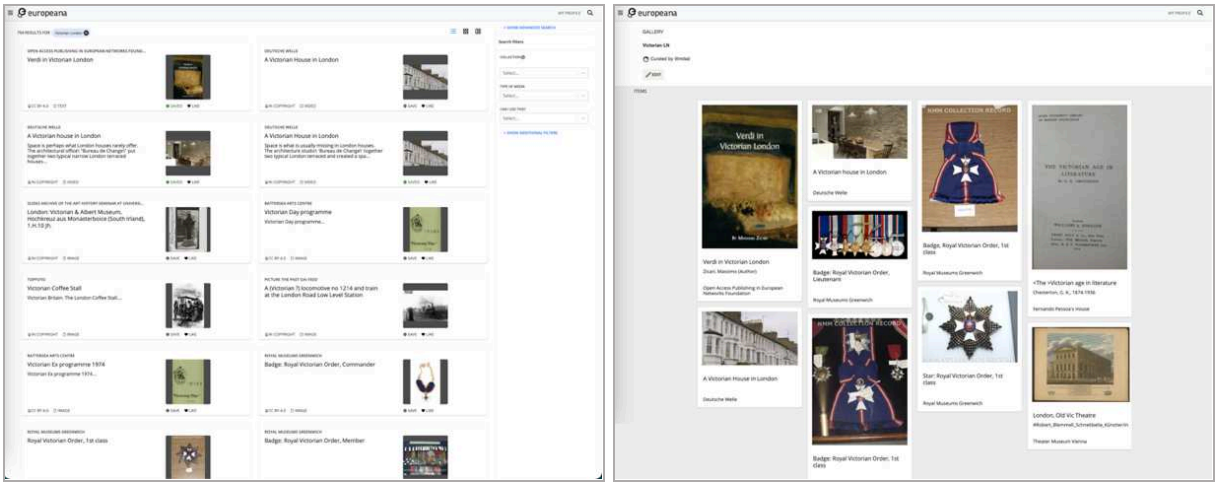

Figure 1: Baseline search interface (left) and workspace (right), closely modelled after the existing Europeana search interface and workspace.

Result-focused

The Result-focused approach keeps the keywords for each search result explicitly linked to the source search result through spatial proximity within a closed region (i.e., providing the keyword list to the right of the search result within the search result card). Each keyword is preceded with an unfilled closed disc, indicating their default interactive status (unselected). Hovering the mouse pointer over the disc or the keyword adds an underline effect and transforms the pointer into a finger icon, providing a visual cue to the interactive nature of the keywords.

Clicking on a keyword adds a unique colour to the disc and the keyword, drawn from a set of ten visually distinct colours generated with ColorBrewer (Harrower & Brewer, 2003). If the selected keyword appears in other search results, those discs and keywords are updated to match this same colour encoding. As a result, the search results that use common keywords become visually linked through the selection of keywords of interest and common colour encoding among the search results. A visual scan reveals the relationships between the search results, as per the Gestalt Principle of Similarity (Koffka, 1935).

This same approach for representing keywords associated with the search results, and making them interactive to reveal relationships is replicated in the workspace, but with a minor change. As the workspace in the baseline approach uses a four-column grid layout, the keywords are listed below the search result, yet still within the search result card. Any keyword selections made during the search process persist when viewing the workspace, and vice versa. Figure 2 provides screenshots of the Result-focused interface, showing both the search results page and the workspace, with a small set of keywords selected.

SERP-focused

The SERP-focused approach aggregates the keywords for all search results on the current search engine results page, providing them in a single list to the right of the search results and above the query refinement filters. The keywords are ordered by frequency, and represented using the same mechanism as in the Result-focused approach (i.e., an unfilled closed disc representing the default unselected state and hover-based underlining to provide a cue to the interactive aspect of the keyword information). These keywords provide an overview of the search results on the current SERP.

As the keywords are disconnected from their source search results, a new approach is needed to reveal the relationships between selected keywords and the search results. An obvious method is to draw a line from each selected keyword to their source search results, but this does not scale well when multiple keywords are selected. Instead, we extend the colour-based similarity approach used in the Result-focused approach, but with some minor modifications. The same colour-encoding mechanism is used when a keyword is selected, resulting in a corresponding colour-filled disc being added at the bottom of the result card for each search result using that keyword. The order and spatial position of the discs are preserved when multiple selections are made, enabling the comparison of search results to one another based on the selected keywords and the searcher's ability to interpret relationships via the same Gestalt Principle of Similarity (Koffka, 1935) noted for the Result-focused approach.

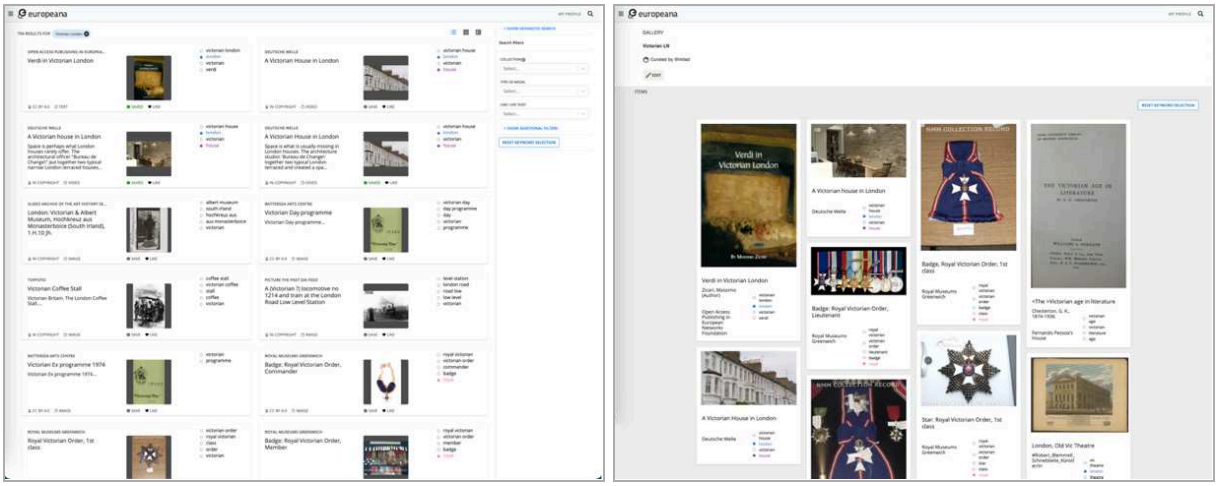

Figure 2: Result-focused keyword-enhanced search interface (left) and workspace (right), providing keywords as a list beside each search result and allowing searchers to select keywords to highlight those same keywords in other search results using colour encoded discs beside each keyword.

This SERP-focused approach is also used in the workspace, where the keyword list is generated only from the search results saved therein. Any selections of keywords made between the search results interface and the workspace are preserved, allowing the searcher to identify aspects of the search results in which they are interested, and to use this to highlight both what has already been found (in the workspace) and what is being sought (in the search results). Figure 3 provides screenshots of the SERP-focused interface, showing both the search results page and the workspace, with a small set of keywords selected.

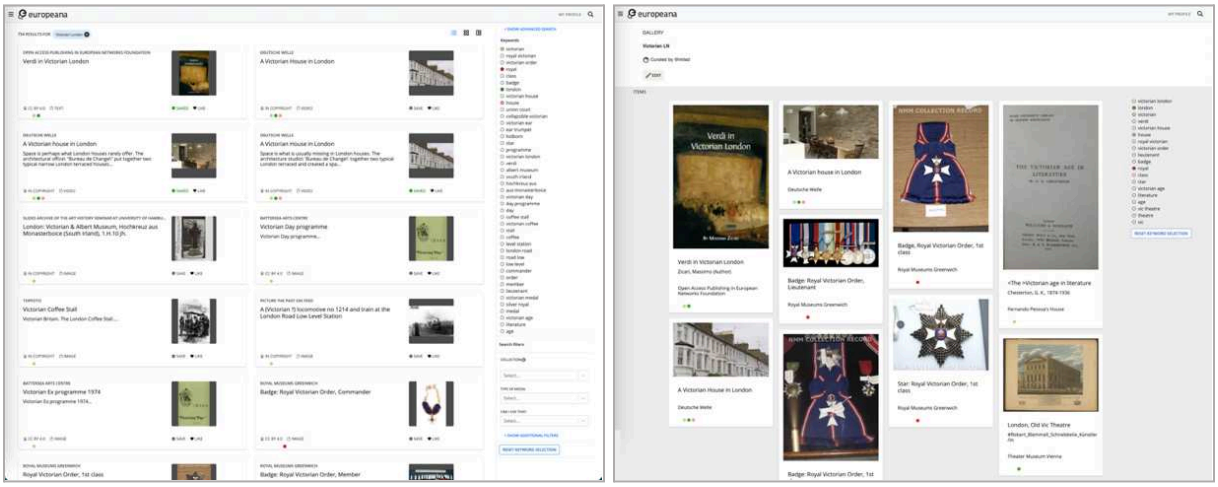

Figure 3: SERP-focused keyword-enhanced search interface (left) and workspace (right), providing keywords in a single list for all search results and allowing searchers to select keywords to mark their source result using colour encoded discs under each search result.

Keyword Extraction

A critical complication in providing keyword-based enhancements to the digital humanities collections such as Europeana is the sparseness of the metadata associated with each search result. The visual keyword/result linking approach originated in an academic digital library search context, where keyword metadata is readily available. Studying this approach in the context of digital humanities collections requires that keywords be inferred from the limited metadata that is present.

We analyzed the metadata that is available and chose the title and description fields as the best source for keyword extraction. This text was combined and provided to an unsupervised machine learning approach for keyword extraction. YAKE! was selected for this purpose, as it has been shown to be effective at extracting keywords from sparse text (Campos et al., 2020).

Empowering Searchers

The visual keyword/result linking approaches used in the two interfaces provide slightly different approaches for empowering searchers with interactive access to information beyond what is provided in the Baseline. Both allow searchers to provide more information to their search process than just the query; they allow searchers to recognize potentially relevant information from the keyword lists, and use this information to support their exploration among the search results through low-effort clicking of potentially relevant keywords. Doing so generates clusters of search results that are visually discernible without changing the content or ranking of the search results. This is a design choice that minimizes the disruption of interactive selections during the search result analysis process.

This process of selecting keywords to show relationships among the search results allows for the examination of transient interests during the exploratory search process. Whether these transient interests are best served by keeping the keywords tightly coupled to the search results (Result-focused) or aggregated over the search results set (SERP-focused) may be influenced by a variety of factors including the nature of the search context and the stage of the exploratory search process.

Methodology

In order to answer RQ2, a controlled laboratory study was designed to assess the impact of manipulating the independent variable (interface type) on a set of dependent variables associated with measuring the value provided to searchers. The three different interface types are those described in the previous section: Baseline, Result-focused, and SERP-focused. A within-subjects design was used, resulting in each participant being exposed to each of the three search interface types while undertaking three different search tasks. A 3x9 Graeco-Latin square was used to vary the order of exposure to the search interface and search task, thus mitigating against order effects.

In this paper, we report on the core of the study, as specified above. The larger study also included having the participants perform a fourth task of their own choosing using their preferred search interface, and included supplemental non-invasive eye-tracking and emotion detection data collection. Results from this data are not reported here due to space limitations.

Tasks

In consultation with a historian familiar with the Europeana collection, three search tasks were developed to create complex information needs that are representative of those typical for digital humanities scholars, with a focus on historical and cultural resources. Following the guidelines of (Kules & Capra, 2009), the tasks were crafted to enable experimental control and were presented as simulated work tasks (Borlund, 2003). Each task was characterized by: (1) an inherent uncertainty necessitating the discovery of new information; (2) a topic that, while unfamiliar, was engaging for the searchers; and (3) detailed descriptions providing adequate context for the searchers to effectively relate to and apply the situation as required. Each task was described in the context of following an exploratory search process that included both an exploratory browsing style of searching and the use of a workspace to review and assess what has been found. The general framework for the task, along with the specific task details are provided in Table 1.

Task Structure (Exploratory Browsing)	Suppose you are starting a research project on [Task Topic]. Your goal is to explore among the Europeana collection for a set of resources that can help you to represent the breadth of this topic. You do not need to perform a deep or comprehensive search right now; instead, your goal is to save a diverse set of resources that will serve as the basis for a more focused search that you will perform later.
Task Structure (Workspace Analysis)	Your goal now is to examine and evaluate the resources you have saved, removing those that are no longer relevant or useful for your task.
Task Topic 1	“the range of Roman artefacts found in England”
Task Topic 2	“how depictions of the New World in maps have changed over time”
Task Topic 3	“how museum collections depict female athletes competing in the Olympics games”

Table 1: Complex search tasks and the associated simulated work task for conducting exploratory searching.

Prior to using the assigned search interface for conducting the assigned search task, a training task was used to showcase the features of the search interface and to provide examples of conducting the initial exploratory browsing searching and the workspace analysis. This included playing a video that showed how the task could be performed using the interface, and then allowing the participant to perform the task on their own. This process, while potentially altering their usual search methods, steered the participants away from employing simple search strategies for complex search tasks. The training task topic was "toys available or popular during the 19th century."

Measures

Two different types of measures were used in this study (utility and perceived value), applied separately to the search interface and the workspace.

The utility measure was divided into two concepts: efficiency and effectiveness. Efficiency was measured by the time it took the participants to complete the task (exploratory browsing or workspace analysis). Effectiveness was measured by the precision of the set of documents found after the exploratory browsing search and the set of documents that remained after the workspace analysis (i.e., the ratio of relevant saved documents to the total number of saved documents).

The perceived value was divided into four concepts: usefulness, ease of use, satisfaction, and user engagement. A modified version of the TAM2 (Venkatesh & Davis, 2000) instrument was used to measure usefulness and ease of use. A set of questions in the same style as the TAM2 instrument was used to measure satisfaction. The UES-SF (O'Brien et al., 2018) instrument was used to measure user engagement. This data was collected via electronic questionnaires using five-point Likert scales, administered at the end of the exploratory browsing search activity and again after the workspace analysis.

Other Data Collection

In addition to measuring data associated with each of the dependent variables, two other data collection instruments were developed. A pre-study questionnaire was used to collect demographic information about each of the participants, including their experience searching within digital humanities collections and use of Euopeana. Pre- and post-task questionnaires were used to measure confidence, interest, and knowledge (on five-point Likert scales) about each search task before starting the search activity and again after completing it.

Study Procedures

The study was conducted in-person at three different universities (two in the United Kingdom, one in Canada) following the procedure outlined in Figure 4. The study took approximately two hours per session. Participants in the UK were compensated £40; participants in Canada were compensated an approximately equal amount of $70. The study was reviewed and approved by the Research Ethics Board at the first author's institution. Confirmation of this review was provided to the counterpart research ethics bodies at the other two institutions (corresponding to the second and last author), where it was also approved.

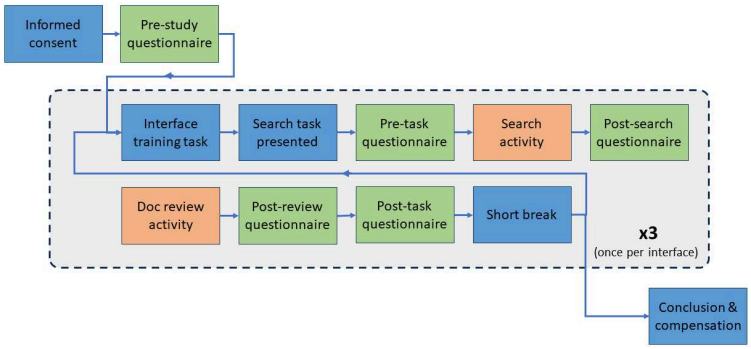

Figure 4: Experimental design

Participants

The recruitment of participants was conducted among digital humanities historians at the authors' three Universities. A snowball sampling method was used, with initial contact made via email to faculty members known to employ digital humanities methods. The recruitment email asked the faculty members to participate in the study, and for distribution of the call for participants to their research staff and graduate students.

Of the 18 participants in the study, one was a senior undergraduate student, five were doctoral students, three were postdoctoral researchers, three were Lecturers/Assistant Professors, four were Senior Lecturers/Associate Professors, and two were Readers/ Professors. Six identified as women; twelve identified as men. None reported having colour vision deficiency. All reported using digital humanities collections at least once a month. One reported "extremely good" search skills with such collections; nine reported "somewhat good"; eight reported "neither good nor bad" search skills. Only two were regular users of Europeana. Although the participants in the study were skewed toward the male gender and had little prior little Europeana use, the other demographic features represent a broad sample of digital humanities historians.

Results

Before and after performing each of the three search tasks, participants were asked to report their degree of confidence, interest, and knowledge on a five-point Likert scale. No statistically significant differences were found between the tasks on these measures. From this, we conclude that the tasks were sufficiently similar to warrant aggregation in all subsequent analysis.

Comparing the aggregated confidence, interest, and knowledge measured before the search task and after, statistically significant differences were found (see Table 2). While no difference was found between the interfaces, this result shows that the participants were actively engaged in the search tasks assigned to them, resulting in significant increases across all three measures.

	Confidence	Interest	Knowledge
Pre-Task	3.11	3.13	2.07
Post-Task	3.43	3.37	2.55
Δ	+ 0.32	+ 0.24	+ 0.48
Wilcoxon SRT	$Z = 75$, $p < 0.05$	$Z = 24$, $p < 0.01$	$Z = 86$, $p < 0.001$

Table 2: Change in confidence, interest, and knowledge as a result of preforming the search tasks.

We now consider whether the manipulation of the independent variable had an impact on the participants' utility and perceived value of both the search interface and the workspace.

Utility of the Search Interface

The time to complete the exploratory browsing search activity and the precision of the saved search results are reported in Table 3. While there were some numerical differences between the mean time to task completion for the participants using each of the interfaces, these were not found to be statistically significant.

Measure	Baseline	Result-focused	SERP-focused	ANOVA
Time to task completion (seconds)	651 (79)	716 (95)	692 (103)	$F = 2.20$, $p = 0.12$
Precision of saved search results	0.66 (0.23)	0.89 (0.12) *	0.85 (0.17) *	$F = 0.841$, $p<0.001$

Table 3: Time to task completion (mean and standard deviation) and precision of the saved search results (* indicates statistically significant differences versus the baseline).

The saved search results were assessed by two members of the research team independently for relevance using a three-point relevance scale (0: not relevant; 1: relevant; and 2: fully relevant). Cohen's kappa showed an inter-coder reliability of $\kappa = 0.897$, indicating strong agreement between the two assessors. Conflicts were resolved via discussions about the relevance of the specific search results. Subsequently, the three-point scale was collapsed into a binary scale, wherein documents deemed partially relevant were categorized as relevant. Based on this, statistically significant differences in the precision of the saved search results were found. Tukey HSD revealed statistically significant differences between the Baseline and both the Result-focused and SERP-focused interface, but no difference between those two interfaces. We conclude that either approach for visual keyword/result linking enables searchers to find and save more relevant documents than when using a traditional search interface that does not provide interactive access to keyword data.

Perceived Value of the Search Interface

The perceived value of the search interface was measured as ease of use, usefulness, satisfaction, and four aspects of user engagement: focused attention, perceived usability, aesthetic appeal, and reward factor. ANOVA results show statistical significance on all measures (see Table 4). Tukey HSD revealed which pairs of conditions contributed to these statistically significant findings. For the ease of use measure, both the Result-focused and SERP-focused interfaces were reported at significantly higher levels than the Baseline, but with no statistically significant difference between them. Similar patterns held for the usefulness and satisfaction measures.

The results were slightly different with respect to the user engagement measures. The SERP-focused interface was reported to provide significantly better focused attention than both the Result-focused and the Baseline interface. For perceived usability and aesthetic appeal, the Result-focused interface was reported at significantly higher levels than the Baseline. For reward factor, the SERP-focused interface was reported as significantly greater than the Baseline.

Measure	Baseline	Result-focused	SERP-focused	ANOVA
Ease of use	3.46	4.13 *	4.17 *	$F = 5.67, p < 0.01$
Usefulness	3.22	3.78 *	3.97 *	$F = 5.87, p < 0.01$
Satisfaction	2.92	3.81 *	3.82 *	$F = 7.88, p < 0.005$
Focused Attention	3.11	3.17	3.96 **	$F = 8.15, p < 0.001$
Perceived Usability	3.11	4.04 *	3.78	$F = 5.48, p < 0.01$
Aesthetic Appeal	2.91	3.67 *	3.50	$F = 4.24, p < 0.05$
Reward Factor	3.69	3.94	4.28 *	$F = 3.44, p < 0.05$

Table 4: Perceived value measures (five-point Likert scale) for all interfaces under investigation (* indicates statistically significant difference versus the Baseline; ** indicates statistically significant difference versus the Baseline and the other visual keyword/document linking approach).

While both of the visual keyword/result linking interface alternatives provided value to the participants, the SERP-focused approach was particularly beneficial in terms of enhancing focused attention and reward factors. Presenting the keywords extracted from each search result as an external list provided an overview of the search results that was difficult to obtain in the other search interfaces (Baseline or Result-focused). This method of displaying and interacting with the keywords enabled the participants to stay focused on the search task and find the activity rewarding. However, with respect to perceived usability and aesthetic appeal, participants preferred the Result-focused approach, where the keywords remained directly associated with their source search results. We conclude that both visual keyword/result linking approaches enhanced the participants' perceived value of the search interface, with the SERP-focused approach being considered superior due to a greater importance being placed on enhancing focused attention and reward factors than perceived usability and aesthetics in the specific context of digital humanities research.

Utility of the Workspace

The time to complete the workspace analysis and the precision of the search results that remain after this analysis are reported in Table 5. Similar to the use of the search interface, there were no significant differences observed in the task completion times between the three conditions.

Measure	Baseline	Result-focused	SERP-focused	ANOVA
Time to task completion (seconds)	119 (24)	139 (23)	132 (35)	$F = 2.38, p = 0.10$
Precision of saved search results	0.79 (0.22)	0.96 (0.05) *	0.90 (0.17)	$F = 5.21, p < 0.01$

Table 5: Time to task completion (mean and standard deviation) and precision of the saved search results (* indicates statistically significant differences).

The set of search results that remained after the participants reviewed what was found in the workspace and removed what was no longer relevant were assessed for precision. For all interfaces, this task resulted in an increase in the precision of the saved search results, showing the overall value of workspaces in supporting exploratory search activities. Statistically significant differences were found between the interfaces; Tukey HSD revealed that the Result-focused approach enabled searchers to prune irrelevant search results from the workspace more effectively than the other two interfaces. We conclude that keeping the keywords directly associated with the search results (Result-focused) is the superior approach for post-search analyses of what has been saved to the workspace.

Perceived Value of the Workspace

The perceived value of the workspace was assessed using the same measures used earlier: ease of use, usefulness, satisfaction, and four aspects of user engagement, all collected after using the workspace. ANOVA results show statistical significance on all but one measure (see Table 6). Tukey HSD revealed which pairs of conditions contributed to statistically significant differences. As with the use of the search interface, after using the workspace participants reported statistically significant positive ease of use responses for the Result-focused and SERP-focused interfaces compared to the Baseline, but with no statistically significant difference between these interfaces. Similar patterns held for the usefulness and satisfaction measures.

For the user engagement measures, more differences between the interfaces were revealed. Both visual keyword/result linking interfaces were reported to have significantly better perceived usability than the baseline. We attribute this to the extra functionality added to these interfaces with the addition of the keyword information, and

the simple interaction mechanisms for revealing relationships. The differences in aesthetic appeal were not found to be statistically significant. Participants reported higher scores for focused attention and reward factor after using the SERP-focused interface compared to the Baseline; there was no statistically significant difference compared to the Result-focused interface. We conclude that both visual keyword/result linking approaches enhanced the participants' perceived value of the workspace. As with the search activity reported earlier, when analyzing what has been found in the workspace, the SERP-focused approach is considered superior due to the benefits it provides for enabling the searcher to stay focused and find the activity being performed rewarding.

Measure	Baseline	Result-focused	SERP-focused	ANOVA
Ease of use	3.69	4.44 *	4.46 *	$F = 6.31, p < 0.005$
Usefulness	3.42	4.10 *	4.15 *	$F = 5.01, p < 0.05$
Satisfaction	3.18	3.94 *	4.00 *	$F = 5.61, p < 0.01$
Focused Attention	2.89	3.19	3.76 *	$F = 4.19, p < 0.05$
Perceived Usability	3.63	4.48 *	4.37 *	$F = 9.33, p < 0.001$
Aesthetic Appeal	3.00	3.57	3.5	$F = 2.26, p = 0.114$
Reward Factor	3.50	3.93	4.32 *	$F = 7.2, p < 0.005$

Table 6: Perceived value measures (five-point Likert scale) for all interfaces under investigation (* indicates statistically significant difference versus the Baseline; ** indicates statistically significant difference versus the Baseline and the other visual keyword/document linking approach).

Conclusion

Many have identified the difficulties that digital humanities scholars face when searching for resources to support their research (Buchanan et al., 2005; Crymble, 2016; Gibbs & Owens, 2012; Hill et al., 2016; Sinn & Soares, 2014; Toms & O'Brien, 2008; Trace & Karadkar, 2017), including those that have explicitly called for the development of better search tools (Wu et al., 2024). As a first step in providing better support to such searchers, we have applied techniques that have been shown to work well in academic digital library search contexts (citation withheld to preserve anonymity during peer review) to the context of searching within digital humanities collections. Doing so provides a new baseline for enabling exploratory search within such collection, against which to compare future interactive information retrieval advancements.

A user study conducted in a controlled laboratory setting revealed slightly different results than what was found in academic digital library searching (citation withheld). In particular, the SERP-focused approach was found to be an improvement over the other approaches with respect to two aspects of user engagement: enabling focused attention and providing reward factors. While both new approaches enable searchers to find initial search results sets that had a higher precision than the baseline, when analyzing what was found in the workspace participants were able to make more effective use of the Result-focused approach for removing irrelevant search results. These mixed results suggest that an approach that works well for one aspect of exploratory search (e.g., exploratory browsing style of searching) may not be the most effective approach for another aspect (e.g., workspace analysis).

These results provide design implications for future work. In particular, separate consideration should be given for how to support activities associated with evaluating new resources found in the search process versus re-evaluation what has been found previously and saved in the workspace. The fundamental distinction between these two cases is whether the searcher is considering something that is unknown (search results) or known (saved results). In this particular study, the evaluation of the workspace occurred immediately after searching; things may change if there is a gap between when the searcher saved the results and when they re-evaluated them in the workspace that is sufficiently long to result in not remembering what is there. Such cross-session search activities have been studied in other contexts (Gomes et al., 2022; Li et al., 2020), and should be studies in this digital humanities context as well.

While these results provide mixed evidence between the two new approaches, they do showcase the value of using keyword data to enable the interactive identification of relationships among search results, both within the search results list during the exploratory browsing style searching and within the workspace during the analysis of what was found. There are opportunities for further study of approaches that reveal relationships among search results, and how such approaches might enable exploratory search processes within digital humanities collections. Such studies may continue to use keywords or leverage other metadata that is of particular interest to digital humanities scholars for showing such relationships (e.g., theme, subject, place, media type, date/era, source).

GENERATIVE AI USE

We confirm that we did not use generative AI tools/services to author this submission.

REFERENCES

Berry, D. M. (Ed.). (2012). *Understanding Digital Humanities*. Palgrave Macmillan UK. https://doi.org/10.1057/9780230371934

Borlund, P. (2003). The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research*, *8*(3). https://informationr.net/ir/8-3/paper152.html

Buchanan, G., Cunningham, S. J., Blandford, A., Rimmer, J., & Warwick, C. (2005). Information Seeking by Humanities Scholars. *Proceedings of the European Conference on Research and Advanced Technology for Digital Libraries*, 218–229. https://doi.org/10.1007/11551362_20

Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., & Jatowt, A. (2020). YAKE! Keyword extraction from single documents using multiple local features. *Information Sciences*, *509*, 257–289. https://doi.org/10.1016/j.ins.2019.09.013

Crymble, A. (2016). Digital library search preferences amongst historians and genealogists: British History Online user survey. *Digital Humanities Quarterly*, *10*(4), Article 4. http://digitalhumanities.org:8081/dhq/vol/10/4/000270/000270.html

Europeana. (2024a, January 5). *Discover Europe's digital cultural heritage*. Europeana. https://www.europeana.eu/en

Europeana. (2024b, January 5). *Documentation on The Europeana Search API*. Europeana PRO. https://pro.europeana.eu/page/search

Gibbs, F., & Owens, T. (2012). Building better digital humanities tools: Toward broader audiences and user-centered designs. *Digital Humanities Quarterly*, *6*(2), Article 2. https://www.digitalhumanities.org/dhq/vol/6/2/000136/000136.html

Gomes, S., Boon, M., & Hoeber, O. (2022). A study of cross-session cross-device search within an academic digital library. *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 384–394. https://doi.org/10.1145/3477495.3531929

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, *40*(1), 27–37. https://doi.org/10.1179/000870403235002042

Hill, T., Charles, V., Stiller, J., & Isaac, A. (2016). "Searching for inspiration": User needs and search architecture in Europeana collections. *Proceedings of the Association for Information Science and Technology*, *53*(1), 1–7. https://doi.org/10.1002/pra2.2016.14505301043

Jackson, A., Lin, J., Milligan, I., & Ruest, N. (2016). Desiderata for exploratory search interfaces to web archives in support of scholarly activities. *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, 103–106. https://doi.org/10.1145/2910896.2910912

Koffka, K. (1935). *Principles of Gestalt Psychology*. Routledge.

Kules, B., & Capra, R. (2009). Designing exploratory search tasks for user studies of information seeking support systems. *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, 419–420. https://doi.org/10.1145/1555400.1555492

Li, Y., Capra, R., & Zhang, Y. (2020). Everyday cross-session search: How and why do people search across multiple sessions? *Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval*, 163–172. https://doi.org/10.1145/3343413.3377970

Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, *49*(4), 41–46. https://doi.org/10.1145/1121949.1121979

Maxwell, D., & Hauff, C. (2021). LogUI: Contemporary logging infrastructure for web-based experiments. *Proceedings of the European Conference on Information Retrieval*, 525–530. https://doi.org/10.1007/978-3-030-72240-1_59

Medlar, A., Ilves, K., Wang, P., Buntine, W., & Glowacka, D. (2016). PULP: A system for exploratory search of scientific literature. *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1133–1136. https://doi.org/10.1145/2911451.2911455

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, *112*, 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

Qvarfordt, P., Tretter, S., Golovchinsky, G., & Dunnigan, T. (2014). SearchPanel: Framing complex search needs. *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, 495–504. https://doi.org/10.1145/2600428.2609620

Sinn, D., & Soares, N. (2014). Historians' use of digital archival collections: The web, historical scholarship, and archival research. *Journal of the Association for Information Science and Technology*, *65*(9), Article 9. https://doi.org/10.1002/asi.23091

Soufan, A., Ruthven, I., & Azzopardi, L. (2022). Searching the literature: An analysis of an exploratory search task. *Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval*, 146–157. https://doi.org/10.1145/3498366.3505818

Toms, E. G., & O'Brien, H. L. (2008). Understanding the information and communication technology needs of the e‐humanist. *Journal of Documentation*, *64*(1), Article 1. https://doi.org/10.1108/00220410810844178

Trace, C. B., & Karadkar, U. P. (2017). Information management in the humanities: Scholarly processes, tools, and the construction of personal collections. *Journal of the Association for Information Science and Technology*, *68*(2), 491–507. https://doi.org/10.1002/asi.23678

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204.

Ward, A. R., & Capra, R. (2021). OrgBox: Supporting cognitive and metacognitive activities during exploratory search. *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2570–2574. https://doi.org/10.1145/3404835.3462790

White, R. W., & Roth, R. A. (2009). Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, *1*(1), 1–98. https://doi.org/10.2200/S00174ED1V01Y200901ICR003

Wu, I.-C., Vakkari, P., & Huang, B.-X. (2024). An exploration of search-as-learning in digital archives of an online museum. *Journal of Documentation*, *80*(2), 298–319. https://doi.org/10.1108/JD-03-2023-0056