Visualization-Enhanced Aggregated Search Interfaces

Milad Momeni*
miladmomeni@uregina.ca
Department of Computer Science
University of Regina
Regina, SK, Canada

ABSTRACT

Search interfaces serve as the primary gateway to information retrieval (IR) platforms, with each IR platform possessing its unique interface. Yet, the challenge remains in delivering diverse content effectively to users, particularly when dealing with unfamiliar sources or content. Instead of platforms presenting search results from different sources in individual tabs for distinct sources, which can lead to user confusion and over-reliance on certain tabs, this paper investigates an approach to aggregate search results into a single, unified display, providing users with a consolidated list. Our research emphasizes novel presentation methodologies that combine insights from previous studies with advanced visualization techniques. The aim is to offer an intuitive and streamlined search experience for all users. Key research questions address: the design of interfaces to blend aggregated results while visually indicating their provenance; the advantages of such interfaces; the impact of search result diversity on perceived trustworthiness; and the applicability of the approach in structured data domains, specifically digital humanities archives and digital academic libraries. Our structured, iterative research methodology encompasses a three-phased approach: starting with low-fidelity prototyping, moving to medium-fidelity design iteration, and currently working towards functional prototype development. An upcoming controlled laboratory study, complemented by data collection tools like LogUI and eye-tracking, aims to gain a comprehensive understanding of user interaction and attention patterns and evaluate the proposed designs, providing insights into their effectiveness and the implications of visual result provenance.

CCS CONCEPTS

• **Information systems** → **Combination, fusion and federated search**; **Rank aggregation**; **Digital libraries and archives**; **Presentation of retrieval results**; **Search interfaces**; • **Human-centered computing** → **Visualization design and evaluation methods**.

KEYWORDS

Aggregated search, information visualization, multi-source search environment, interface design and evaluation, controlled laboratory study

CHIIR '24, March 10–14, 2024, Sheffield, United Kingdom

ACM ISBN 979-8-4007-0434-5/24/03.
https://doi.org/10.1145/3627508.3638336

ACM Reference Format:
Milad Momeni. 2024. Visualization-Enhanced Aggregated Search Interfaces. In *Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '24), March 10–14, 2024, Sheffield, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3627508.3638336

1 INTRODUCTION

The domain of information retrieval now encompasses many information-seeking tasks, ranging from image and video searches to scholarly articles and real-time social media updates. This diversification has led to platforms tailored to support interactions with a broad range of internet content. A challenge for these platforms is ensuring the efficient delivery of varied content to users, especially when the sources of content might be unfamiliar. Aggregation of search results into a comprehensive list integrating various sources has been a traditional approach to this challenge [7, 14].

This study focuses on a platform that provides an interface for users to classify and comment on a broad range of online content. Search results are presented in individual tabs, with each tab representing a distinct content source. This setup assumes that users approach their search with a definitive purpose and awareness of where to find answers. The segmented nature of this structure presents several usability issues: (1) searchers who are not invested in conducting a thorough search will rely on the default tab as the sources of the search results; (2) potential misdirection resulting from selecting an unsuitable tab; and (3) the risk of users abandoning their search due to unsatisfactory findings [10]. Tabbed interfaces like this segment the importance of transitioning towards a more integrated display of search results.

To address this, there is a growing inclination towards aggregating search results into a unified list. This approach encompasses the merger of varied sources to enhance selection, and showcasing the combined results in a manner that facilitates user appraisal. Research has delved into the intricacies of aggregated search, pinpointing effective algorithms for source inclusion and result blending [2]. Numerous studies have analyzed the impact of mixed search results on user efficacy and perceptions [1, 2]. It has been suggested that the speed at which an individual perceives can influence the usability of blended interfaces [30]. This further emphasizes the need to design and assess aggregation methodologies adaptable to different processing speeds.

Meeting these challenges requires a multi-faceted approach. Central to this is the methodology of information representation. Current academic investigations underline the advantages of emphasizing distinct content sources and types to amplify user engagement [5, 10]. Findings on blended search interfaces provide insights into user behaviour during interactions with consolidated Search Engine Results Pages (SERPs) [20], the interplay of varied content sources,

and the ties between cognitive capabilities and search outcomes [3, 30]. Through the application of information visualization techniques, users can discern patterns and focus on important details, avoiding the need for reading everything [15]. Improving the way results are displayed in search platforms is a research priority.

This research seeks to bridge existing knowledge gaps by exploring innovative presentation methodologies in platforms supporting global content interactions. By fusing insights from prior investigations and harnessing advanced visualization techniques, the aim is to redefine user interactions with search outcomes. The overarching ambition is to craft a smooth, intuitive, and efficient search journey for users, when their information needs require content from multiple sources.

2 RESEARCH QUESTIONS

In pursuit of how to blend and present search results from different sources within interfaces, while ensuring the provenance of these results remains transparent, we recognized the necessity to explore the intrinsic advantages and broader applicability of these interfaces. This led us to formulate the following research questions:

RQ1: How can an interface be designed to blend aggregated results from various sources while visually representing provenance in the search interfaces?

RQ2: What are the advantages and potential disadvantages of designing interfaces that blend aggregated results from various sources while visually representing provenance?

RQ3: How does manipulating the diversity of search results in an aggregated search interface impact users' perception of the trustworthiness of the results?

RQ4: How can this approach be applied to different structured data domains, such as digital humanities archives and digital academic libraries?

3 RESEARCH METHODOLOGY

3.1 Approach

To address the challenges and complexities of designing interfaces that aggregate search results from varied sources while visually signifying their origins, we embarked on a structured and iterative methodological approach.

In the initial phase of *Alternative Solution Ideation, Planning, and Low-Fidelity Prototyping*, we tackled the complexities of multi-source search environments inherent to our platform. This stage was pivotal for devising potential solutions, grounded in recent scholarly work [14] and diverse methodologies. We augmented our strategy by integrating principles from visual perception and cognitive psychology, enhancing users' ability to visualize the origin of search results and discern patterns or connections in the data. Leveraging theories like pre-attentive processing, opponent process theory of colour, Mackinlay's APT, Bertin's semiology of graphics, and Gestalt principles [8, 21, 29], we anchored our approach in established information visualization tenets.

During this phase, we sketched multiple low-fidelity prototype interfaces, each representing a distinct approach to search results aggregation. Post-development, these prototypes underwent heuristic evaluations, guided by the methodologies of Norman and Molich [24]. This involved applying specific heuristics such as consistency, feedback, error prevention, and flexibility. The evaluations were instrumental in refining the prototypes, ensuring they were intuitive and met user needs.

Our approach crystallized into three primary methods. First, the **Streamlined Fusion** strategy, which amalgamates search results into a unified list. Secondly, the **Dynamic Result Tiles**, a method we scrutinized for its limitations. Despite certain drawbacks—such as potential information overload and category overlap—we adopted it for its unique ability to segregate search results into distinct, topic-specific tiles. This method was particularly useful in scenarios where distinct categorization of information was paramount. Lastly, the **Keyword Grouping** technique visually clusters search results around pertinent keywords or tags, highlighting their relevance to specific topics or themes. Despite the identified limitations of the Dynamic Result Tiles, its inclusion alongside the other two methods provided a comprehensive approach, catering to a diverse range of user preferences and search contexts.

Progressing to the *Medium-Fidelity Prototype Development, Formative Assessment, and Design Iteration* stage, our prototypes were crafted to blend aggregated data while presenting their origins. Here, we took forward all three of the low-fidelity prototypes and created medium-fidelity prototypes in Figma (sketch/design software). Following this, we proceeded to conduct cognitive walkthroughs [9] of the medium-fidelity prototypes. This analysis allowed us to further refine them, addressing any identified issues. Preliminary evaluations rooted in user-centred design principles [25] ensured iterative and collaborative refinements in conjunction with user feedback to our prototypes, keeping our designs in line with user needs and project goals.

During the *Functional Prototype Development Phase*, we are in the process of system integration, where we are integrating multiple systems to provide a cohesive user experience. After evaluating the medium-fidelity prototypes, we are finalizing which designs to develop into functional software. We employed the MERN Stack for development [23]. For data interactions, we integrated with an external GraphQL API. This integration streamlined our data queries and reduced over-fetching. Using Apollo Server capabilities from the external API, we efficiently handled queries and ensured consistent data retrieval for our users.

For comprehensive user understanding, we'll utilize both behavioural and attentional data. User interaction patterns will be recorded using LogUI [22], providing invaluable insights into navigational habits and preferences. Complementing this, to gain a deeper perspective on user attention and to understand how users are interacting with the visual elements added to search results, we incorporated the Neon eye-tracking glasses from Pupil Labs [28]. This state-of-the-art technology enabled us to discern areas of maximum focus, offering an unfiltered lens into the visual prominence of aggregated search results.

Additionally, to provide a structured approach toward interface evaluation, we'll adopt Wilson's taxonomy [32] for search interface features. This classification, encompassing input, control, informational, and personalization categories, serves as an analytical compass in our assessment.

3.2 Experimental Design

In line with the pursuit of understanding the intrinsic advantages and broader applicability of search result aggregation and presentation methods, a user study is envisioned. The planned structure is outlined as follows:

For **RQ1**, which explores the innovative design of interfaces to integrate aggregated results from multiple sources and visualize their provenance, we have conceptualized and are poised to implement four unique interfaces: the baseline tab-based search interface and three novel designs — Streamlined Fusion, Keyword Grouping, and Dynamic Result Tiles. The sole *independent variable* in this study pertains to the methodology used to represent and interact with the aggregated search results within the interface.

RQ2 focuses on the evaluation of the advantages and implications of visual presentation of search result provenance. The foundation rests upon a controlled laboratory study [19].

Earlier studies emphasized methodologies comparing a predicted aggregated SERP to an ideal reference for specific queries [4]. Another approach introduced a utility-centric metric, evaluating dimensions such as the relevance of verticals on the SERP and the cognitive effort users expend on them [34]. These methodologies underwent validation through user-centric studies [2].

Building on foundational research, our primary dependent variables include Efficiency, Effectiveness, Feature use, User Engagement, User Satisfaction, Usefulness, and Ease-of-use. Regarding Usability, Usefulness examines if search features meet participant needs [14], measured by a post-task questionnaire using four TAM2 questions [31]. Ease-of-use assesses interface intuitiveness [14], evaluated similarly to usefulness. User Engagement reflects user investment with the interface, measured post-task using twelve UES short-form questions [26]. Satisfaction gauges user-derived positive emotions post-interaction, determined via a four-question survey [16, 33]. The nature of the data associated with these dependent variables underscores the need for statistical analysis methods.

Our approach also categorizes interaction based on advanced features, referencing Wilson's taxonomy [32]. We utilize the LogUI JavaScript framework for real-time interaction logging [22]. Effectiveness and Efficiency are analyzed in terms of selected search results' precision and time taken for task completion, respectively, within the interface.

Participants begin with a Qualtrics survey explaining the study. Demographics and prior knowledge are collected via pre-study and pre-task questionnaires. Each task includes video and written training. Real-time logging tracks task time, viewed documents, and interface use. Post-task questionnaires gather usability feedback similar to pre-task.

Pre-task and post-task questionnaires, real-time logging mechanisms using LogUI, and eye-tracking are the principal tools employed for data collection. Given the inherent challenges such as participant fatigue and potential biases, measures like short breaks have been incorporated. Recognizing the potential for individual differences in search behaviours or domain knowledge to influence interactions, these nuances will be factored in during data analysis. Our protocols have been carefully devised, and we are in the process of submitting an REB application to the University's Research Ethics Board to ensure ethical standards are upheld.

By measuring heatmaps, saccades, fixations, and dwell time, specialized glasses for eye tracking enable User Behaviour Analysis. Heatmaps aggregate gaze data to depict areas most frequently viewed. Saccades capture the rapid eye movements across an interface, fixations reveal areas where users pause to process information, and dwell time quantifies the duration users engage with specific elements [11, 13, 17, 18].

To delve into **RQ3**, we design a controlled experiment manipulating the diversity of search results, grounded on the principle of the *Observation of Diversity in Search*. We recognize the importance of presenting a breadth of information, encompassing diverse perspectives from various sources, for comprehensive user understanding. In tandem with this, the *Contribution to Trustworthiness* serves as a vital foundation. By visually showcasing the origin of each search result, we aspire to support user trust and encourage critical evaluations of content reliability.

We'll conduct a controlled experiment where participants engage with our search interface under two conditions: High Diversity (a wide range of perspectives from multiple sources) and Low Diversity (results mainly from a narrow set of sources).

Participants will be given specific queries to interact with under both conditions. The sequence of conditions will be counterbalanced to offset order effects [27]. Post-interaction, participants will rate the trustworthiness of results using established Likert scale questionnaires [6]. Eye-tracking metrics will further provide insights into user attention and behaviour [17].

Content quality will remain consistent across both conditions, ensuring that perceived trustworthiness differences arise from diversity, not content credibility [12]. Using our visualization-enhanced aggregated search interface, we aim to understand the relationship between search result diversity and trustworthiness.

For **RQ4**, we aim to extend the scope of our investigation to examine the feasibility of implementing our aggregated search interfaces across varied domains, especially those marked by structured data. Notably, our focus will be on digital humanities archives and digital academic libraries, both rich repositories of structured data supported by intricate metadata frameworks. While harnessing this data is crucial, it's of paramount importance to maintain the innate contextual significance of these datasets. This not only ensures data integrity but also facilitates user searches with satisfactory findings. By drawing from the insights garnered from RQ1 and RQ2, this research question aims to adapt and fine-tune our search interfaces to best serve these specific, structured data realms.

4 FUTURE PLANS

Building on our work that explores innovative methodologies for search result aggregation and presentation, our next steps are structured around our four main research questions. For **RQ1**, we aim to finalize the development and testing of our aggregated search interfaces, emphasizing visual representation of provenance. This will involve iterative improvements and refinements based on user feedback. Under **RQ2**, we plan to conduct a comprehensive laboratory study to evaluate the advantages and implications of our visual presentation methodologies. Emphasis will be placed on usability metrics, including user engagement, satisfaction, and efficiency. To address **RQ3**, we are designing a controlled experiment that

manipulates the diversity of search results, exploring the potential impact on users' perceptions of trustworthiness. This will involve a comparative analysis between high-diversity and low-diversity result presentations, measuring trustworthiness through established questionnaires and eye-tracking metrics. Finally, for **RQ4**, our goal is to extend our research to digital humanities archives and digital academic libraries. We will examine the adaptability of our search interfaces to structured data domains, ensuring data integrity and relevance. In tandem with these objectives, we are in the process of obtaining ethical approval for our planned studies, ensuring that all research activities align with ethical guidelines.

5 EXPECTED CONTRIBUTION

Our research focuses on improving search result aggregation and presentation. We will introduce a user-centric visual representation of provenance, backed by empirical evidence from comprehensive studies. By exploring the impact of result diversity on trustworthiness, we'll offer insights into optimizing user trust. Extending our innovations to digital humanities archives and academic libraries will showcase adaptability across structured data domains. Collectively, our contributions aim to offer meaningful enhancements to visualization-enhanced aggregated search interfaces.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Orland Hoeber, for his invaluable guidance throughout this project. We appreciate the support from Mitacs through the Accelerate Grant and Scrawlr Development Inc. as our partner.

REFERENCES

[1] Jaime Arguello. 2012. Federated Search in Heterogeneous Environments. *SIGIR Forum* 46, 1 (may 2012), 78–79. https://doi.org/10.1145/2215676.2215686

[2] Jaime Arguello. 2017. Aggregated Search. *Foundations and Trends® in Information Retrieval* 10, 5 (2017), 365–502. https://doi.org/10.1561/1500000052

[3] Jaime Arguello and Rob Capra. 2016. The Effects of Aggregated Search Coherence on Search Behavior. *ACM Trans. Inf. Syst.* 35, 1, Article 2 (sep 2016), 30 pages. https://doi.org/10.1145/2935747

[4] Jaime Arguello, Fernando Diaz, Jamie Callan, and Ben Carterette. 2011. A Methodology for Evaluating Aggregated Search Results. In *Proceedings of the 33rd European Conference on Advances in Information Retrieval* (Dublin, Ireland) *(ECIR'11)*. Springer-Verlag, Berlin, Heidelberg, 141–152. https://doi.org/10.1007/978-3-642-20161-5_15

[5] Jaime Arguello, Wan-Ching Wu, Diane Kelly, and Ashlee Edwards. 2012. Task Complexity, Vertical Display and User Interaction in Aggregated Search. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Portland, Oregon, USA) *(SIGIR '12)*. Association for Computing Machinery, New York, NY, USA, 435–444. https://doi.org/10.1145/2348283.2348343

[6] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123. https://cir.nii.ac.jp/crid/1370857593794409373

[7] Nicholas J. Belkin. 2016. People, interacting with information. *SIGIR Forum* 49, 2, 13–27. https://doi.org/10.1145/2888422.2888424

[8] Jacques Bertin. 1983. *Semiology of Graphics*. University of Wisconsin Press. https://doi.org/10.5555/1095597

[9] Marilyn Hughes Blackmon, Peter G. Polson, Muneo Kitajima, and Clayton Lewis. 2002. Cognitive Walkthrough for the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 463–470. https://doi.org/10.1145/503376.503459

[10] Horatiu Bota, Ke Zhou, and Joemon M. Jose. 2016. Playing Your Cards Right: The Effect of Entity Cards on Search Behaviour and Workload. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval* (Carrboro, North Carolina, USA) *(CHIIR '16)*. Association for Computing Machinery, New York, NY, USA, 131–140. https://doi.org/10.1145/2854946.2854967

[11] Andrew Duchowski. 2007. *Eye Tracking Techniques*. Springer London, London, 51–59. https://doi.org/10.1007/978-1-84628-609-4_5

[12] Andrew J. Flanagin and Miriam J. Metzger. 2000. Perceptions of Internet Information Credibility. *Journalism & Mass Communication Quarterly* 77, 3 (2000), 515–540. https://doi.org/10.1177/107769900007700304 arXiv:https://doi.org/10.1177/107769900007700304

[13] Joseph H Goldberg and Xerxes P Kotval. 1999. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics* 24, 6 (1999), 631–645. https://doi.org/10.1016/S0169-8141(98)00068-7

[14] Marti Hearst. 2009. *Search User Interfaces*. Cambridge University Press. https://doi.org/10.5555/1631268

[15] Orland Hoeber. 2018. Information Visualization for Interactive Information Retrieval. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval* (New Brunswick, NJ, USA) *(CHIIR '18)*. Association for Computing Machinery, New York, NY, USA, 371–374. https://doi.org/10.1145/3176349.3176898

[16] Orland Hoeber and Soumya Shukla. 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 (2022), 1171–1191. https://doi.org/10.1002/asi.24623 arXiv:https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/asi.24623

[17] Kenneth Holmqvist and Richard Andersson. 2017. *Eye-tracking: A comprehensive guide to methods, paradigms and measures*. 746 pages.

[18] Robert J.K. Jacob and Keith S. Karn. 2003. Commentary on Section 4 - Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. In *The Mind's Eye*, J. Hyön, R. Radach, and H. Deubel (Eds.). North-Holland, Amsterdam, 573–605. https://doi.org/10.1016/B978-044451020-4/50031-1

[19] Diane Kelly. 2009. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends® in Information Retrieval* 3, 1—2 (Jan. 2009), 1–224. https://doi.org/10.1561/1500000012

[20] Zeyang Liu, Yiqun Liu, Ke Zhou, Min Zhang, and Shaoping Ma. 2015. Influence of Vertical Result in Web Search Examination. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Santiago, Chile) *(SIGIR '15)*. Association for Computing Machinery, New York, NY, USA, 193–202. https://doi.org/10.1145/2766462.2767714

[21] Jock Mackinlay. 1986. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.* 5, 2 (apr 1986), 110–141. https://doi.org/10.1145/22949.22950

[22] David Maxwell and Claudia Hauff. 2021. LogUI: Contemporary Logging Infrastructure for Web-Based Experiments. In *Advances in Information Retrieval*, Djoerd Hiemstra, Marie-Francine Moens, Josiane Mothe, Raffaele Perego, Martin Potthast, and Fabrizio Sebastiani (Eds.). Springer International Publishing, Cham, 525–530. https://doi.org/10.1007/978-3-030-72240-1_59

[23] MongoDB. 2023. *How To Use MERN Stack: A Complete Guide*. https://www.mongodb.com/languages/mern-stack-tutorial

[24] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press, New York, NY, 249–256. https://doi.org/10.1145/97243.97281

[25] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.

[26] Heather L. O'Brien, Paul Cairns, and Mark Hall. 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 (2018), 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

[27] Harold Pashler. 1997. *The Psychology of Attention*. The MIT Press. https://doi.org/10.7551/mitpress/5677.001.0001

[28] PupilLabs. 2023. *Neon & Cloud Documentation*. https://docs.pupil-labs.com/neon

[29] Dejan Todorovic. 2008. Gestalt principles. *Scholarpedia* 3, 12 (2008), 5345.

[30] Lauren Turpin, Diane Kelly, and Jaime Arguello. 2016. To Blend or Not to Blend? Perceptual Speed, Visual Memory and Aggregated Search *(SIGIR '16)*. Association for Computing Machinery, New York, NY, USA, 1021–1024. https://doi.org/10.1145/2911451.2914739

[31] Viswanath Venkatesh and Fred D. Davis. 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Manage. Sci.* 46, 2 (feb 2000), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

[32] Max Wilson. 2011. *Search-User Interface Design*. Morgan & Claypool Publishers. https://doi.org/10.2200/S00371ED1V01Y201111ICR020

[33] Hong (Iris) Xie and Dietmar Wolfram. 2002. State digital library usability: Contributing organizational factors. *Journal of the American Society for Information Science and Technology* 53, 13 (2002), 1085–1097. https://doi.org/10.1002/asi.10152 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.10152

[34] Ke Zhou, Ronan Cummins, Mounia Lalmas, and Joemon M. Jose. 2012. Evaluating Aggregated Search Pages. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Portland, Oregon, USA) *(SIGIR '12)*. Association for Computing Machinery, New York, NY, USA, 115–124. https://doi.org/10.1145/2348283.2348302