

ARGO SCHOLAR: Interactive Visual Exploration of Literature in Browsers

Kevin Li*, Haoyang Yang*, Anish Upadhayay*, Zhiyan Zhou*, Jon Saad-Falcon*, Duen Horng (Polo) Chau*

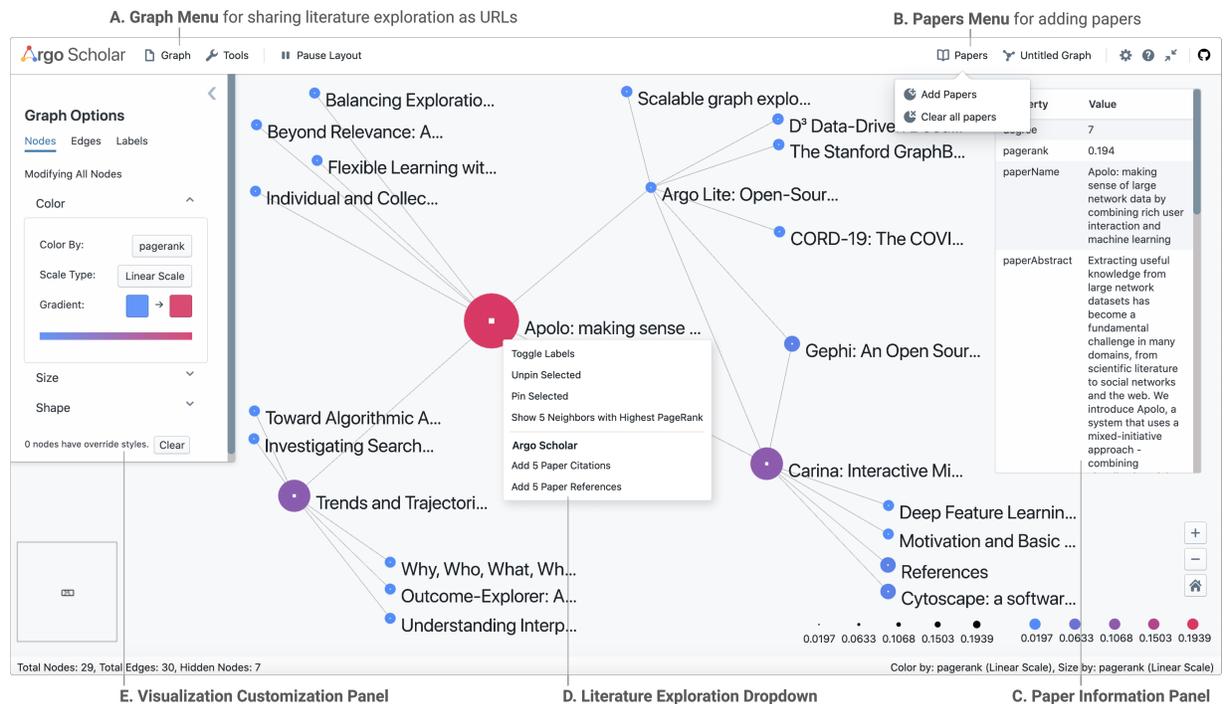


Figure 1: ARGO SCHOLAR visualizing a literature network of papers (nodes) and citations (edges) around the *Apolo* article. **A. Graph Menu** for saving and sharing user’s literature exploration snapshots as URLs. **B. Papers Menu** for creating new exploration and adding papers via their Semantic Scholar CorpusIDs. **C. Paper Information Panel** shows key information (e.g., paper title, abstract, citation count). **D. Literature Exploration Dropdown** for incrementally adding more connected papers that are citing or being cited. **E. Visualization Customization Panel** for modifying visual properties of the graph of papers (e.g., color, size, labeling).

ABSTRACT

Discovering and making sense of relevant research literature is fundamental to becoming knowledgeable in any scientific discipline. Visualization can aid this process; however, existing tools’ adoption and impact have often been constrained, such as by their reliance on small curated paper datasets that quickly become outdated or a lack of support for personalized exploration. We introduce ARGO SCHOLAR, an open-source, web-based visualization tool for interactive exploration of literature and easy sharing of exploration results. ARGO SCHOLAR queries and visualizes Semantic Scholar’s live data of almost 200 million papers, enabling users to generate personalized literature exploration results in real-time through flexible, incremental exploration, a common and effective method for researchers to discover relevant work. Our tool allows users to easily share their literature exploration results as a URL or web-embedded IFrame application. ARGO SCHOLAR is open-sourced and available at <https://poloclub.github.io/argo-scholar/>.

*Georgia Institute of Technology.

{kevin.li,alexanderyang,aupadhayay3,zzhou406,jonsaadfalcon,polo}@gatech.edu

Index Terms: Human-centered computing—Visualization—Visualization systems and tools;

1 INTRODUCTION

Exploring and making sense of relevant research literature is fundamental to any scientific discipline. While visualization tools have been proposed to aid this critical task, their adoption and impact have often been constrained by their reliance on manually-curated paper datasets that were often small and outdated subsets of the available literature [1, 2]. A recent visual tool called Connected Papers aims to automate the discovery of relevant work for users, but it shows all users the same visualization for a given paper and does not support incremental exploration or associative browsing, which are common effective strategies for discovering relevant work [1, 3, 4] (e.g., starting with a handful familiar or influential papers and branching out over time). Addressing the common literature exploration needs shared by researchers, our ongoing work makes the following contributions:

1. **ARGO SCHOLAR, an open-source, web-based visualization tool for interactive exploration of literature and easy sharing of exploration results** (Fig. 1). Anyone can access ARGO SCHOLAR using their web browser on desktops and mobile devices without the need for any software installation. ARGO SCHOLAR is open-sourced under the permis-

sive MIT license¹ and is available at the public link: <https://poloclub.github.io/argo-scholar/>.

- Incremental, flexible literature exploration with up-to-date data.** ARGO SCHOLAR connects to the live data of Semantic Scholar [5] via its API, gaining access to its 194 million indexed papers (as of June 2021). ARGO SCHOLAR’s ability to query and visualize up-to-date data in real-time overcomes major limitations faced by prior tools proposed for literature sensemaking that relied on manually-curated paper datasets that were often small and outdated subsets of available literature [1, 2]. Users can add papers indexed by Semantic Scholar by entering the papers’ unique CorpusIDs (Fig. 1B). And for each paper, users can incrementally explore the connected papers that cite it or are cited, generating personalized literature networks for each user (Fig. 1D), unlike existing tools such as Connected Papers² that show all users the same visualization for a given paper.
- Easy sharing of literature exploration results.** ARGO SCHOLAR allows users to share their literature networks and visualization as a URL or IFrame web-embedded application (Fig. 1A). By accessing a shared link, researchers build upon existing exploration progress to discover more relevant papers.

2 ARGO SCHOLAR’S SYSTEM DESIGN

Developed using modern web technologies, ARGO SCHOLAR’s interface provides users with new functionalities beneficial for literature sensemaking (e.g., incremental exploration, live data connection to Semantic Scholar) while inheriting the core graph visualization capabilities of Argo Lite [6] (e.g., React for Blueprint for UI, MobX state management, Three.js WebGL graph rendering).

Displaying Paper Relations. ARGO SCHOLAR’s main view displays a citation network as a graph of papers (nodes) and citations (edges). Users can add any paper indexed by Semantic Scholar [5] (<https://www.semanticscholar.org>) by entering the paper’s unique CorpusID via the *Papers Menu* (Fig. 1B). From there, the user can branch out from currently visualized papers by exploring their citations and references or input new, potentially unrelated, papers. Directed edges point from a citing paper to the cited; edge directions can be toggled. A variety of visual properties of the nodes, edges, and node labels can be customized via the *Customization Panel* (Fig. 1E). For example, nodes can be colored and sized based on paper attributes such as citation count, with the ranges of both being user-adjustable.

Incremental Exploration. Making sense of literature through incremental paper exploration or associative browsing is a common, effective method for researchers to discover relevant work [1, 2]. Specifically, they may start with a handful of (familiar) papers, then gradually build their network of related papers through exploring the references and citations of their growing network. ARGO SCHOLAR supports this literature sensemaking strategy; users can iteratively add 5 reference or citation papers for a paper³ via its *Exploration Dropdown* (Fig 1D). The paper is connected to the added papers by directed edges that encode the citation relationships previously described, and the added papers are aligned vertically to improve label readability. Network data, such as the degree, pagerank, size, and color of existing nodes, are updated accordingly.

Learning More About Each Paper. The *Paper Information Panel* (Fig 1C) of ARGO SCHOLAR shows essential paper information for a highlighted or hovered-over node, e.g., paper title, abstract,

authors, citation count, publication venue, publication year, and the URL to its Semantic Scholar page (for paper PDF, figures, etc.).

Saving and Sharing Explorations Across Platforms. ARGO SCHOLAR allows users to save and share their literature exploration as a *snapshot* JSON file that stores the information of all explored papers, their connections, and visualization customizations via the *Graph Menu* (Fig 1A). Users have two options when saving and sharing their exploration snapshots: locally on their device or in the cloud. If users opt for the first option, they will download the JSON file and then upload it to ARGO SCHOLAR whenever they wish to resume their sensemaking process. Or, users can save their snapshots to ARGO SCHOLAR’s server for free, enabling them to share their literature network as a custom URL, bypassing the need to send JSON files. Saving the snapshot to ARGO SCHOLAR’s servers also grants the ability to embed the literature network as either an HTML IFrame or a Jupyter Notebook IFrame. ARGO SCHOLAR runs on all modern web browsers (e.g., Chrome, Safari, Firefox, Edge), broadening its access.

3 CONCLUSION AND ONGOING WORK.

ARGO SCHOLAR is an open-source, web-based visualization tool for interactive exploration of literature and easy sharing of exploration results. ARGO SCHOLAR queries and visualizes Semantic Scholar’s live data, enabling users to generate personalized literature exploration results in real-time through flexible, incremental exploration. Anyone can access our tool using their web browser on desktops and mobile devices without the need for any software installation.

We plan to improve the user experience of ARGO SCHOLAR by enabling users to directly search and show paper results in ARGO SCHOLAR, bypassing the current need for locating the paper’s CorpusID. This tighter connection with Semantic Scholar would allow us to enhance the usability and user-friendliness of our visualization tool. Additionally, we plan to improve the ordering of adding cited papers and referenced papers by implementing the ability to add them based on metrics such as citation count, recency, or relevance. These improvements would add more flexibility to the sensemaking process as users can further customize the way they explore literature. We also plan to evaluate ARGO SCHOLAR through user studies with both beginning and seasoned researchers to evaluate ARGO SCHOLAR’s usability and its ability to help with literature sensemaking in the long term. As ARGO SCHOLAR improves its functionalities, we look forward to more students, researchers, and practitioners adopting ARGO SCHOLAR to discover relevant research and make sense of the literature across disciplines.

REFERENCES

- [1] S. Kairam, N. Riche, S. Drucker, R. Fernandez, and J. Heer. Refinery: Visual exploration of large, heterogeneous networks through associative browsing. *Computer Graphics Forum*, 34, 2015.
- [2] Duen Horng Chau, Aniket Kittur, J. Hong, and C. Faloutsos. Apolo: making sense of large network data by combining rich user interaction and machine learning. *CHI*, 2011.
- [3] Peter Pirolli and Stuart Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, volume 5, pages 2–4. McLean, VA, USA, 2005.
- [4] Robert S. Pienta, J. Abello, Minsuk Kahng, and Duen Horng Chau. Scalable graph exploration and visualization: Sensemaking challenges and opportunities. *BIGCOMP*, pages 271–278, 2015.
- [5] Kyle Lo, Lucy Lu Wang, Mark Neumann, Rodney Michael Kinney, and Daniel S. Weld. S2orc: The semantic scholar open research corpus. In *ACL*, 2020.
- [6] Siwei Li, Zhiyan Zhou, Anish Upadhayay, Omar Shaikh, Scott Freitas, H. Park, Zijie Jay Wang, Susanta Routray, Matthew Hull, and Duen Horng Chau. Argo lite: Open-source interactive graph exploration and visualization in browsers. *CIKM*, 2020.

¹<https://github.com/poloclub/argo-scholar>

²<https://www.connectedpapers.com>

³Semantic Scholar supports 100 requests per 5 minutes per IP address