International Society of the Learning Sciences

AI in Service of Epistemic Agency: Interactive Visualizations for Knowledge Building Communities

Lydia Cao, Marlene Scardamalia, Emily Liu
ly.cao@utoronto.ca, marlene.scardamalia@utoronto.ca, emilyxiaofei.liu@mail.utoronto.ca
University of Toronto

Abstract: Knowledge Building communities seek to advance community knowledge and push the frontier of their field through sustained and collaborative effort. While AI has been increasingly used to support researchers to keep pace with the rapid evolution of the field and navigate large volumes of publications, existing AI tools often produce finished products with minimal human involvement. This tech demonstration presents the Knowledge Building Space Navigator (KBSN), which foregrounds epistemic agency by generating AI-powered interactive visualizations that enable researchers to explore knowledge landscapes from multiple perspectives. Using Knowledge Building Summer Institute (KBSI) proceedings (2020-2024) as a case study, we applied BERTopic modeling, semantic density mapping, and social network analysis to establish a proof of concept for KBSN, which aims to help researchers identify research trends, knowledge gaps, opportunities for innovation, and potential collaboration. This prototype aims to demonstrate a new model for integrating AI into knowledge creation processes that foreground epistemic agency to support knowledge advancement.

Introduction

Knowledge building communities or knowledge-creating organizations, such as research institutions, educational organizations, professional networks, and innovation hubs, strive to advance the state of the art in their field (Scardamalia & Bereiter, 2021). Knowledge Building, with its emphasis on progressive discourse and idea improvement, demands mechanisms for communities to continuously assess and leverage their evolving community knowledge to advance ideas (Scardamalia, 2002). While traditional synthesis methods like systematic reviews (for research communities, for example) provide valuable and critical insights, they often cannot keep pace with the rapid evolution of community knowledge. Annual research conference proceedings and abstracts (such as those from the ISLS annual conference) represent the very cutting-edge development in scholarly fields. Yet little has been done to synthesize these contributions largely due to the sheer volume. As a result, we often lack the bird's-eye view to better understand where our field currently stands, identify emerging trends, and recognize knowledge gaps. This represents a significant missed opportunity for collective knowledge advancement.

To transform knowledge repositories into actionable resources for knowledge advancement, we need both better mechanisms for agile scoping of community knowledge and collective idea improvement. This technical demonstration presents the early-stage development of a prototype, namely, Knowledge Building Space Navigator (KBSN), as an attempt to tackle this challenge.

AI-powered tool that foregrounds epistemic agency

There are a number of existing AI-powered research tools (e.g., Elicit, Consensus, Scite) that help researchers to identify, summarize, and verify the credibility of research articles. However, they tend to generate summaries/synthesis as finished products without necessarily involving humans in the process of co-construction. As a result, they serve as automated producers of intellectual products rather than partners with shared cognitive responsibility. Generating finished products could potentially compromise epistemic agency that characterizes authentic knowledge creation, that requires humans to negotiate meaning, judge areas of improvement and promisingness of ideas (B. Chen et al., 2015). Some AI-powered tools (e.g., Connected papers and Rabbit research) produce networks and give researchers more agency to navigate a knowledge space, but are primarily based on citation and co-authorship. Researchers, however, often need to understand other conceptual affinities beyond bibliometric relationships. Overall, AI-powered research tools have been designed to produce *summative results* rather than *conceptual artifacts* that can be leveraged to improve ideas (Bereiter, 2002).

Knowledge Building Space Navigator (KBSN) is fundamentally designed as a Knowledge Building tool that fosters idea improvement and collaboration. KBSN produces interactive network and dynamic visualizations as *conceptual artifacts* that allow users to examine a knowledge terrain from various vantage points of interest and granularity. Some of the features of KBSN in development include: 1) topic modeling that reveals the evolution of research topics with Large Language Models (LLMs), allowing users to adjust parameters such as cluster numbers to generate visualizations of various granularities; 2) semantic density mapping to identify research gaps and opportunities for rise-above; 3) social network analysis to visualize existing collaborations and identify potential collaborators; 4) multiple vantage points toggle/filter capabilities across dimensions like research topics, theoretical frameworks, methodologies, and geographic locations.

Knowledge Building Summer Institute (KBSI) as a Case Study

To better illustrate the features and functions of KBSN, in this tech demo, we used Knowledge Building Summer Institute (KBSI) proceedings spanning the years 2020-2024 as a proof of concept. KBSI is an international annual gathering dedicated to advancing knowledge building theory, pedagogy, and practice (Knowledge Building International, n.d.). Since its inception in 1997, KBSI has served as a vital hub where researchers, educators, practitioners, engineers, designers, and policymakers from around the world come together to share their most recent work, discuss challenges, and collaboratively advance the field. However, the distributed nature of this expanding global community and network presents significant challenges to effectively coordinate effort and form opportunistic collaboration across institutional and geographical boundaries (Y. Chen & Zhang, 2024).

The need for a tool like KBSN emerged during the organization of KBSI 2025, where we need to better understand the evolution of KBSI over past years, assess our progress to date, and identify research partnerships (both established collaborations and potential collaborations between researchers unaware of each other's work). This prototype represents an early step toward developing a user-friendly tool that could support better coordination and collaboration in a large international community for collective knowledge advancement.

Method

We analyzed 143 research abstracts from KBSI proceedings spanning 2020-2024. Our approach combined advanced natural language processing, large language models, network analysis, and various visualization techniques. The goal of the section is to outline the computational approach as a proof of concept –validation and benchmarking will be conducted in subsequent experiments.

Topic Modeling

We used BERTopic, which leverages transformer-based language models for better semantic relationship processing compared to traditional methods like Latent Dirichlet Allocation (LDA) (Grootendorst, 2022; McInnes et al., 2020). UMAP is used for dimensionality reduction and HDBSCAN (Campello et al., 2013) for clustering papers into research topics. We then used a large language model, in this case, OpenAI's GPT4o with structured prompting with three abstracts from each cluster to generate interpretable topic labels.

Semantic Density Map

In addition to Topic Modelling that categorizes papers into discrete categories, we created a Semantic Density Map that represents similarity relationships among papers (Börner et al., 2003; Qiu & Miikkulainen, 2024). We used UMAP for dimensionality reduction and Gaussian Kernel Density Estimation (KDS) to quantify the concentration of papers in different regions of the semantic space This density measure indicates how well explored a certain research area is in KBSI. Furthermore, we identified the following categories of papers:

1) **Topic-Edge Papers**: Papers at the periphery of a topic (greater than two standard deviations from their topic centroid).
2) **Cross-Topic Papers:** Papers that have strong semantic connections to multiple research topics (having at least three nearest neighbors from different topic clusters).
3) **Boundary Papers:** Papers with similar distance to multiple research topics sitting at the boundary between topics (the ratio of distances to the closest and second-closest topic centroids approaching 1).

Social Network Analysis

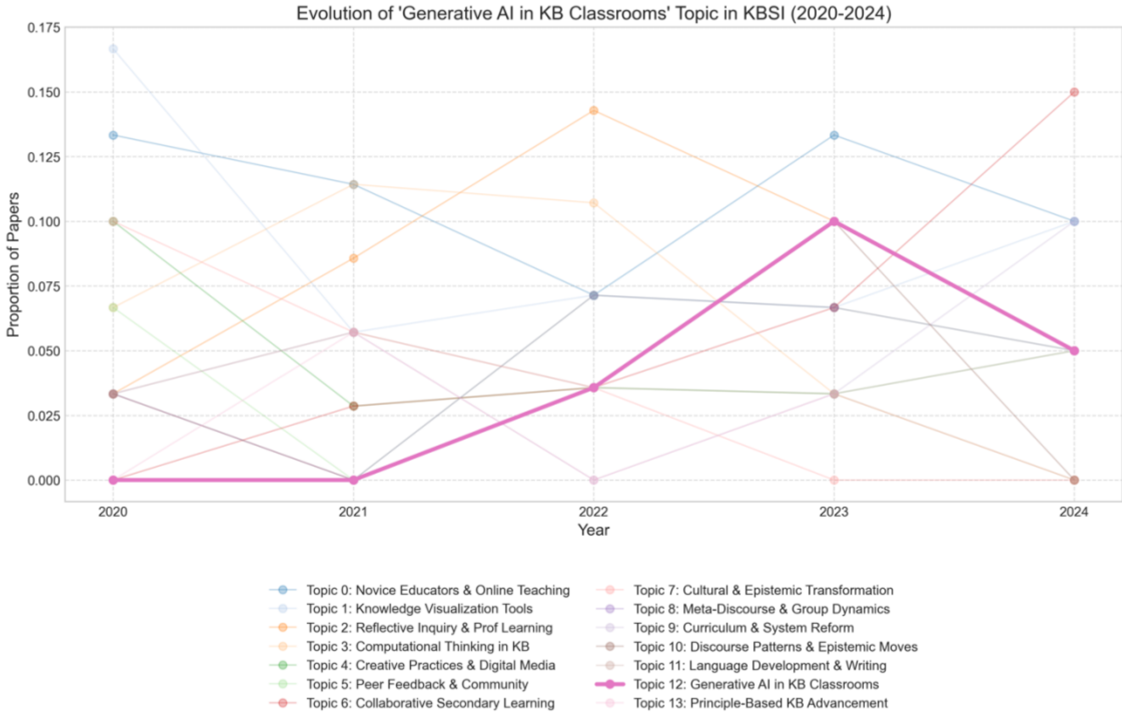
We constructed two social networks to visualize research collaboration patterns (Wasserman & Faust, 1994). The Existing Collaboration Network represents established collaborations, with nodes representing individual authors (attributed by primary research topic) and weighted edges indicating frequency of co-authorship. To identify potential collaboration opportunities, we developed a separate network that connects researchers working on common topics who haven't yet worked together. The edges between two authors indicate that they share at least one research focus but have not co-authored any papers, which suggests potential future partnerships. Both networks were visualized using force-directed layouts with color-coding by research topic.

Findings

Evolution of Research Trends

It is important to note that the following visualizations are intended to be interactive and dynamic, allowing users to adjust parameters (e.g., cluster number) to have control over the granularity of the view and to highlight a particular area/topic/paper for closer examination. Thus, the results should not be interpreted as definitive, but rather *conceptual artifacts* to be used for idea advancement. Figure 1 shows 14 distinct thematic clusters/research topics in KBSI over the past 5 years, with Topic 12, Generative AI in KB classrooms highlighted as an example.

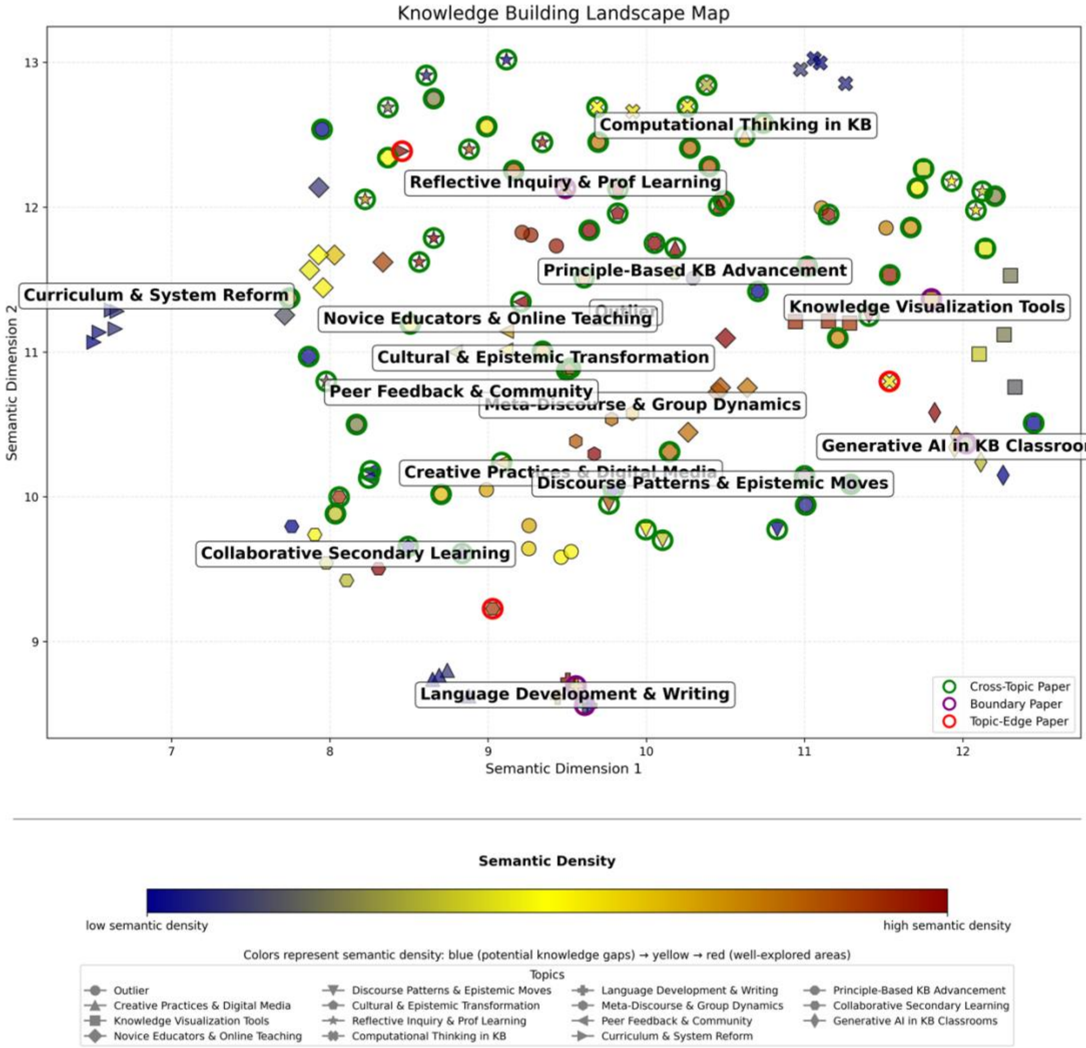
Figure 1 Evolution of Research Topics in KBSI (2020-2024)



Semantic density map to identify potential research gaps and opportunities

Figure 2 shows the Semantic Density Map for KBSI 2020-2024. Each node represents a paper, with the shape corresponding to its research topic. The color of each node indicates semantic density. Warmer colors (e.g., red) represents high density, indicating well-researched areas. Colder colors (e.g., blue) represents low density, indicating potentially under-researched areas that might represent knowledge gaps or emerging research frontiers. The further apart nodes appear on the map, the more semantically distant they are from each other. Papers with high density values are located in regions with many other similar articles, while those with low density values are positioned in relative isolation. To facilitate easier navigation of the knowledge building landscape, we also organized papers into several categories. It is important to note that these are interpretations rather than definitive classifications. Researchers can start with a bird's-eye view of the knowledge landscape, then gradually focus on specific areas and papers of interest.

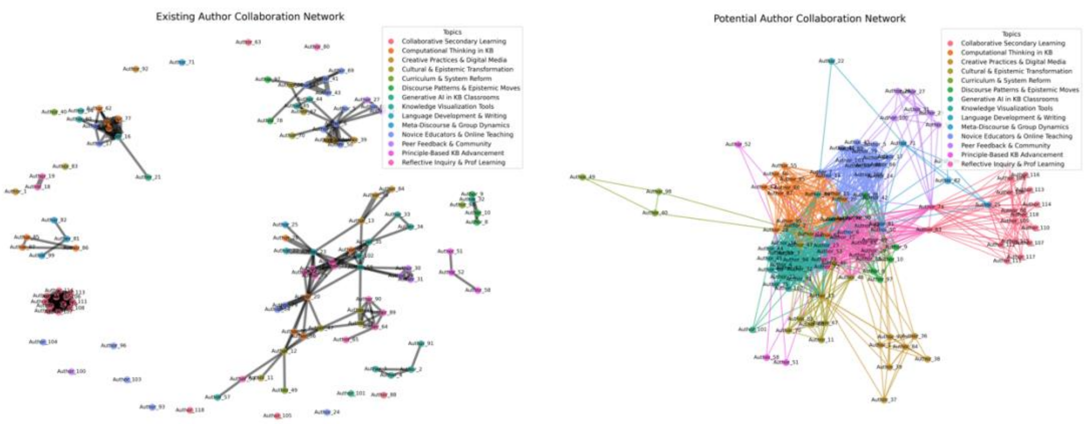
1) **Topic-Edge Papers** (3 papers, 2.1%, highlighted in red circle): Papers at the periphery of their corresponding topic clusters, potentially extending domain boundaries in novel directions.
2) **Cross-Topic Papers** (75 papers, 52.4%, highlighted in green circle): Papers with connections to multiple research topics, indicating cross-topic integration. The high percentage of cross-topic papers (52.4%) suggests significant cross-topic research across the knowledge building community.
3) **Boundary Papers** (15 papers, 10.5%, highlighted in purple circle): Papers with a similar distance to multiple topic centroids, potentially suggesting an opportunity for rise-above, i.e., increasing explanatory power and going beyond existing frameworks. We found that 19.4% of low-density papers are also boundary papers, suggesting that some knowledge gaps and innovations exist at the boundaries between established research areas.



Social network among researchers

Figure 3 shows two research networks: the left is based on existing collaborations (co-authorships) and the right is based on potential collaborations (researchers working in the same research area but without co-authored work). This interactive social network is designed to help researchers identify potential collaborators.

Figure 3 Research Topics and Associated Researchers



Conclusion and Next Steps

KBSN represents a new model of how AI can be integrated into knowledge creation processes. Rather than positioning AI as an automated producer of intellectual products, KBSN demonstrates how AI can serve epistemic agency by being a partner with shared cognitive responsibility. Through this approach, KBSN has the potential to transform research communities' ability to continuously assess and leverage their evolving collective knowledge to accelerate knowledge advancement.

Our goal for this tech demo is to solicit a wide range of feedback on the early-stage conception and visualization approaches of KBSN from the learning sciences and knowledge building community. The community's insights and feedback will be integrated to shape and refine the design of KBSN. For example, we are planning to include collaborative annotation to transform these visualizations into improvable community conceptual artifacts. After this redesign phase, we will conduct validation experiments and benchmarking with larger datasets that incorporate broader knowledge building literature to further improve the computational methods. We will then implement user testing to evaluate usability and effectiveness. The ultimate goal of KBSN is to enhance collaboration and accelerate knowledge advancement for broader research and knowledge building communities across disciplines.

References

Bereiter, C. (2002). *Education and mind in the Knowledge Age* (pp. xiii, 526). Lawrence Erlbaum Associates Publishers.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, *37*(1), 179–255. https://doi.org/10.1002/aris.1440370106

Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-Based Clustering Based on Hierarchical Density Estimates. In J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 160–172). Springer. https://doi.org/10.1007/978-3-642-37456-2_14

Chen, B., Scardamalia, M., & Bereiter, C. (2015). Advancing knowledge-building discourse through judgments of promising ideas. *International Journal of Computer-Supported Collaborative Learning*, *10*(4), 345–366. https://doi.org/10.1007/s11412-015-9225-z

Chen, Y., & Zhang, Y. (2024). The use of distance-shortening strategies to enhance opportunistic collaboration in knowledge-building environments. *Educational Technology Research and Development*, *72*(6), 3159–3184. https://doi.org/10.1007/s11423-024-10389-3

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure* (No. arXiv:2203.05794). arXiv. https://doi.org/10.48550/arXiv.2203.05794

Knowledge Building International. (n.d.). *Knowledge Building Summer Institutes*. Retrieved January 17, 2025, from https://ikit.org/kbi/index.php/summer-institutes/

McInnes, L., Healy, J., & Melville, J. (2020). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction* (No. arXiv:1802.03426). arXiv. https://doi.org/10.48550/arXiv.1802.03426

Qiu, X., & Miikkulainen, R. (2024). *Semantic Density: Uncertainty Quantification for Large Language Models through Confidence Measurement in Semantic Space* (No. arXiv:2405.13845). arXiv. https://doi.org/10.48550/arXiv.2405.13845

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp. 67–98). Open Court.

Scardamalia, M., & Bereiter, C. (2021). Knowledge Building: Advancing the State of Community Knowledge. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 261–279). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_14

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.