BridgeIt: Crossing the Knowledge Gaps Between Maker Communities of Practice

7th International Symposium on Academic Makerspaces

ISAM 2023 Poster No.: 127

Adam G. Emerson[1], Cesar Torres[2]

[1]Adam G. Emerson; Dept. of Computer Science Eng, University of Texas at Arlington; e-mail: adam.emerson@mavs.uta.edu

[2]Dr. Cesar Torres; Dept. of Computer Science Eng, University of Texas at Arlington e-mail: cearto@uta.edu

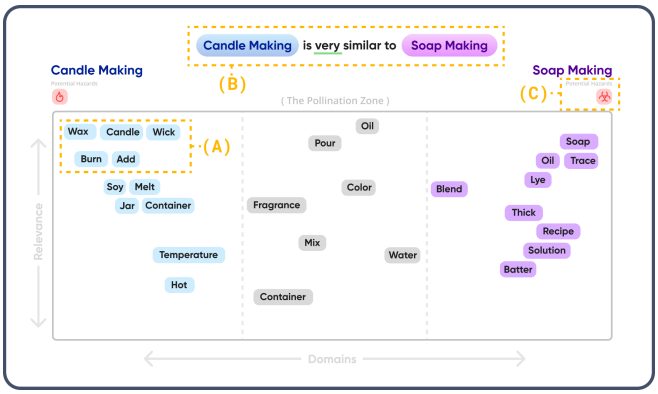

Figure 1:BridgeView, highlighting the similar set of fragrances and oils used by soap-making and candle making; (A) Extracted concepts from the domains, (B) At-a-glance description of domain similarity, (C) Hazard labels indicate potential safety concerns.

Introduction

In recent years, online tutorials on video sharing platforms like TikTok and YouTube [5] have become a prominent resource for DIY and makerspace learning, especially in offering valuable insight on the tools, materials, and techniques used. *However, existing platforms often lack the comprehensive tools needed to navigate and connect knowledge across different DIY domains, creating a barrier for newcomers who must start from scratch when learning new skills.* This result is hours spent watching video tutorials, mostly consisting of redundant information in hopes of finding answers to specific questions. The problem is compounded when operating in interdisciplinary practices because maker communities can be isolated in discrete "filter-bubbles" by search-engine algorithms which partition and segregate information into niche sub-communities whose motivations differ [6].

To overcome this knowledge isolation challenge, we introduce BridgeIt, a knowledge discovery tool. BridgeIt allows both new and experienced makers to quickly understand a Community of Practice (CoP) [2] by visualizing shared concepts between two communities, encouraging users to compare and contrast domains. Leveraging YouTube, BridgeIt collects tutorial transcripts from various communities, extracts key concepts, and presents them in an easily digestible format, supporting the development of a common language between otherwise disparate maker groups. Through effective aggregation and dissemination of community-based knowledge, BridgeIt empowers makers to collaborate and explore new interdisciplinary mediums.

Related Work

Existing video sharing platforms lack comprehensive tools to navigate and connect knowledge across DIY domains. BridgeIt addresses this by presenting an advanced system for cross-domain knowledge discovery, leveraging natural language processing (NLP) and Topic Modeling techniques [3, 4] to analyze video tutorial transcripts and extract domain concepts. The work builds on insights from ConceptScope [1], a system which uses Bubble Treemaps to represent conceptual relationships within specialized research domains from mined text documents. While ConceptScope excels for domain-specific documents, BridgeIt aims to overcome its interdisciplinary limitations by employing a dynamic concept extraction technique as opposed to a rigid ontology-based approach. By treating each community as their own subject-matter expert, BridgeIt provides a more intimate view of the practices there-in. This language-first approach is what makes BridgeIt an effective tool for the cross-pollination of ideas among diverse groups of makers.

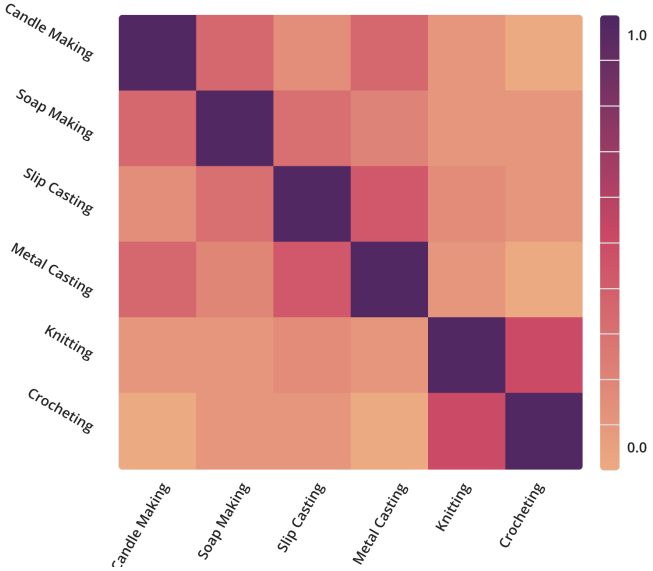

Figure 2: A correlation matrix highlighting the overlap in language (%) between various CoPs. Significant overlap can be seen in the textiles and casting-and-molding sub-groups.

Methods

We used a system building methodology to explore how knowledge fragmentation occurs across different maker communities and to develop effective strategies for bridging those fractures. Currently, BridgeIt supports side-by-side comparison-and-contrast of two maker domains simultaneously. The system has crawled 451 transcripts, across 6 domains (Fig 2.), and enumerated over 906 concepts. The process has been automated such that new domains can be crawled at a rate of 5.5 minutes of video per second, or around 2 minutes per 100 videos. Concepts are extracted via NLP by first selecting all the nouns and verbs in each transcript. This pool of words is checked against a curated "stop-word" list tailored to filter out noise related to video tutorials including common phrases like "Click to subscribe" or "Like and comment". Each concept is assigned a score for each domain using TF-PDF [3], a modified version of TF-IDF that considers concept frequency and proportion of transcripts a concept occurs in.

BridgeIt consists of two core interactions:

- *BridgeView*, a scatter-plot style graph that positions concepts as interactive pills between domains. (Fig. 1).
- The *Concept Pane*, a panel providing users with a more intimate understanding of concepts and their position within a CoP (Fig. 3).

BridgeView

To understand how BridgeIt can be used in practice, we describe a motivating example: consider Luci, a maker with a background in candle making. Luci is exploring soap making for the first time and turns to BridgeIt to get started. After selecting the two domains, Luci is met with the *BridgeView* (Fig. 1), a scatter-plot style chart, filled with concepts from both candle making and soap making. The concepts are positioned along a bi-polar x axis so that the candle making concepts fall to the left, and the soap making concepts to the right. Concepts in the middle of the graph are highlighted as the "bridge words", representing the intersection of the two CoPs and are good opportunities for Luci to begin building up her soap making knowledge. At the top of the *BridgeView* Luci notices safety labels have been highlighted for each domain (Fig. 1). Candle making is labeled as a "Burn Hazard", and soap making has a label for "Chemical Hazard". Hovering over the "Chemical Hazard" label, Luci notes that BridgeIt has highlighted Lye as a potential irritant, a material she does not have experience with in her prior work with candles.

Concept Pane

To get a better understanding of how Lye is used in soap making, Luci clicks on its concept pill. The *Concept Pane* (Fig. 3) expands, and Luci is met with several new views. An AI insight is automatically generated, further explaining the context of lye, stating that it: *"likely refers to a caustic substance that is used in soap making to create a chemical reaction with oils or fats."*

Below the AI insight, a temporal line graph indicates to Luci

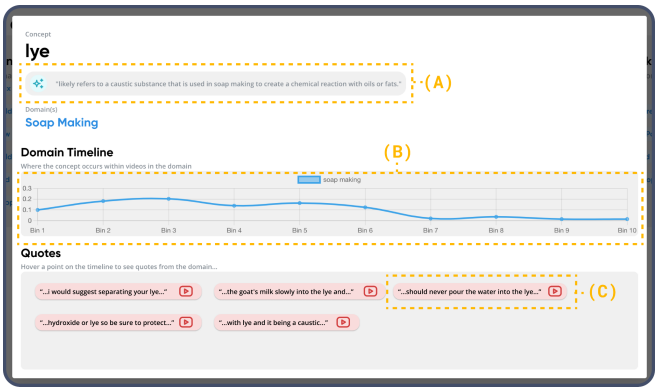

Figure 3: A Concept Pane for "Lye"; (A) The automated AI Insight, (B) Domain timeline indicating frequency of use over tutorial length, (C) Context quotes pulled from video transcripts

when lye is most frequently coming up during soap making tutorials. Luci notes that lye is frequently mentioned towards the beginning and middle of tutorials, but hardly at all towards the end. From this she deduces that lye may be an essential material in the early stages of the soap making process. Luci hovers a point on the line towards the middle of the graph and is met with quotes about lye taken directly from that time position of video tutorials: *"...you should never pour water into Lye..."* one quote reads, warning Luci to take extra precaution. Each quote is accompanied by a hyperlink which directs Luci to the respective video tutorial at the precise time of the highlighted quote. Collectively, these interactions provide Luci with a basis of knowledge on lye and its uses in soap making, ensuring she can proceed with confidence in her soap making journey, and onward to other concepts.

Discussion

While BridgeIt's system for cross-domain knowledge discovery successfully extracts valuable domain concepts, and achieves its goal of bridging gaps between diverse DIY domains, we acknowledge there are several present shortcomings. Not every extracted topic appears equally salient to its domain, indicating room for improvement. We plan to address this by better integrating phrase mining techniques like TopMine [4] that allow us to extract n-grams more effectively. Our vision for BridgeIt includes a robust user tagging system that allows users to further contextualize concepts by superimposing their own perspectives on top of a given domain. This tagging system could act as an individual tool, provide further opportunity for collaboration amongst makers, enabling them to create their own bridges. Our future efforts will continue to focus on the spread of knowledge between CoPs, ensuring that BridgeIt empowers artisans, makers, and enthusiasts with the most relevant and valuable concepts, and fosters a collaborative and inclusive DIY community.

Acknowledgements

This work was supported by the NSF REU Site on Hybrid Design and Fabrication (CNS-2150321).

References

[1] X. Zhang, S. Chandrasegaran, and K.-L. Ma, “Conceptscope: Organizing and visualizing knowledge in documents based on domain ontology,” *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021. doi:10.1145/3411764.3445396

[2] E. Wegner, *Communities of Practice: A Brief Introduction*. 2011.

[3] Khoo Khyou Bun and M. Ishizuka, “Emerging topic tracking system,” *Proceedings Third International Workshop on Advanced Issues of E-Commerce and Web-Based Information Systems. WECWIS 2001*, 2001. doi:10.1109/wecwis.2001.933900

[4] A. El-Kishky, Y. Song, C. Wang, C. R. Voss, and J. Han, “Scalable topical phrase mining from text corpora,” *Proceedings of the VLDB Endowment*, vol. 8, no. 3, pp. 305–316, Nov. 2014. doi:10.14778/2735508.2735519

[5] R. Kaysen, “A new generation of influencers has discovered D.I.Y. on a tiny budget,” The New York Times, https://www.nytimes.com/2022/01/28/realestate/diy-small-budget-influencers.html (accessed Jul. 24, 2023).

[6] L. Parramore and E. Pariser, “The filter bubble,” The Atlantic, https://www.theatlantic.com/daily-dish/archive/2010/10/the-filter-bubble/181427/ (accessed Jul. 24, 2023).