Crowdsourcing Exploratory Cues for Idea Browsing and Inspiration Discovery

Xiaotong (Tone) Xu[1], Khuyen Le[2], Christopher Cha[3], Samuel Fleet[1], Steven P. Dow[1]

[1] Dept of Cognitive Science, University of California San Diego, La Jolla, California 92037
[2] Computer Science Dept, Stanford University, Stanford, California 94305
[3] Dept of Computer Science & Engineering, University of California San Diego, La Jolla, California 92037
xt@ucsd.edu, khuyenle@stanford.edu, ccha@ucsd.edu, sfleet@ucsd.edu, spdow@ucsd.edu

Abstract

Advances in text analysis have made it possible to highlight language associations between ideas that might reveal tacit dimensions of the larger design space and help innovators relate them more effectively. But what kind of metadata should be surfaced to help creative exploration? To explore this, we use a human computation approach to transform online community ideas to multiple semantic representations of ideas with a design problem-solving schema. We crowdsourced metadata related to this schema on 1k+ of product ideas from ProductHunt.com. We then implemented an idea exploration interface, based on the collected product ideas and associated metadata, that recommends similar ideas and surfaces their design-relevant associations. Using this design probe, we conducted a formative study with nine design novices to understand if and what kinds of language associations might facilitate creative exploration and iteration.

To support creative exploration, inspiration seekers have to understand what has been done thoroughly, particularly examples that are similar to their own ideas. One must infer relationships between examples and make comparisons and contrast as they navigate (Gentner and Markman 1997; Chang, Hahn, and Kittur 2020). It also takes strategies to find relevant ideas and to see how they explicitly relate because the real-world ideas are often ill-structured and loosely connected (Lopez-Vega, Tell, and Vanhaverbeke 2016). Research has shown that human annotators can make such data more structured by marking metadata with expert-driven schema (Chan et al. 2018; Hope et al. 2021). Advances in NLP can compare semantic representations of metadata within different dimensions of the schema to find inspirational ideas that are novel yet relevant (Chan et al. 2018; Hope et al. 2017, 2021). The prior work has focused on narrowing down ideas for inspiration but less is known about how surfacing the schematic metadata on examples serve as exploratory cues for users navigate the design space and discover new inspirations. To explore this, we created a creativity-support tool that allowed inspiration seekers to explore the design space related to seeding ideas based on crowdsourced metadata along a design problem-solving schema.

This current work provides three main contributions:

- We present a data preparation pipeline where we adopt a human computation approach to convert real-world ideas produced by an online open community to a dataset that supports fellow innovators for future creative exploration.
- We implement an ideator-centered example exploration interface that leverages the latest NLP techniques and supports real-time navigation of the design space.
- We conduct a user study and qualitative interviews with nine design novices with a basic project-based learning experience to investigate the efficiency of our tool and how language associations may play a role in creative exploration and idea iteration. We summarize our findings into design insights for future example browsers that leverage human computation data.

Dataset Preparation

We scraped 11,699 unique ideas from ProductHunt in reverse chronological order from February 2021 to January 2018. Of these ideas, we sampled a subset of 1.5k ideas based on popularity from productivity and technology categories to fit the context of our formative study. We crowdsourced metadata about them with a schematic questionnaire from a previous study on design problem-solving (Xu, Fan, and Dow 2020). The metadata questionnaire asked crowdworkers to provide information on five dimensions (stakeholder, problem, solution, context, and design goal) for individual ideas. In the crowdsourced data, we used a series of standards to assure the metadata quality, such as avoid abusing and answering the questions correctly. We ended up collecting valid metadata for a total of 1,357 product design ideas using the schema template, yielding a total of 6,785 pieces of metadata.

Using Sentence Transformers's linguistic extractions, we extracted keywords from the metadata. We encoded the text into feature vectors and created a pairwise matrix of cosine similarity scores between the extracted metadata and the collection of extracted metadata of products in the database. From this, we created a list of related ideas for each metadata with the similarity score stored in the dataset for use in the design probe.

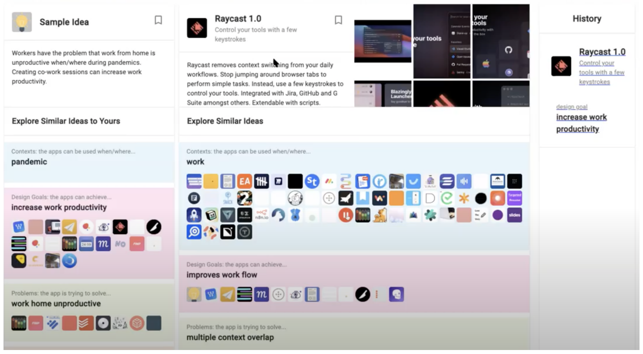

Figure 1: The screenshot of the design probe UI.

Design Probe

Our design probe helps ideators compare their ideas with existing ideas in the ProductHunt database. It consists of a front-end interface, a web server, and a database, implemented with React, Node.js, Express, Flask, and MongoDB. Users can type their ideas in the same schematic format and Flask processes ideas in back-end and fetch back to the main UI. The main UI comprises three distinct vertical components that each serve a different purpose and a common horizontal divide for our two left components. The leftmost component is the seeding idea analysis panel, the central component is the comparison panel, and the right panel is tracking the navigation history.

The idea panel presents a seeding idea on top and the metadata based on dimensions beneath it. Each metadata is accompanied by icons representing similar ideas to the user's own idea, which are fetched from the back-end. The user can hover over each of the icons to see a popover displaying the title and the tagline of the app.

Once the user clicks on an icon on the idea analysis panel, the top part of the comparison panel shows this selected idea. The bottom part shows its metadata associated with each dimension and the apps most similar to those metadata. The user can hover over each app icon to see a preview of that app. This allows the user to browse apps related to their idea by design dimensions. By navigating related apps, the user can also branch out to apps that are two, three, etc. degrees related to their idea.

Preliminary Study

We conducted an online user study, followed by a qualitative interview with nine design students with two goals: (1) To validate whether the human computation approach we adopted could produce robust information that supports the multi-faceted example exploration from the user study. (2) To investigate what kind of descriptive language associations help creative exploration and iteration from the open interview.

Nine participants (six females) were recruited via university email lists who have taken at least one college-level project-based learning course related to web app design or development. All participants had prior experience in brainstorming for technology products.

The participants took remote studies with an experimenter via Zoom. To simplify the brainstorming task and focus on creative exploration, we provide the same cliche seeding idea to all participants at the start of the test. The participants were told that it was an idea produced by a previous ideator, and their task was to learn its similar products, understand the existing design space, and identify room for innovation. They were told to spend at least 15-minute on creative exploration, which they should also think aloud what they navigated or found inspirational.

Formative Results

We summarize our observations from the study and synthesize those into three design insights for building our next iteration of creativity-support tools that leverage human computation approaches.

Design Insight #1: Surface and categorize explicit associations between relevant ideas. According to the participants, our semantic representations of ideas were robust in describing the correlated ideas and recommending relevant ideas. All participants reported that the metadata was descriptive about their corresponding examples, providing relevant information. Because the metadata is associated with relevant product ideas, participants mentioned that the ideas listed under each metadata were similar.

> If we were to just take that original text and then display a whole bunch of related APP ideas in general, I think that this (tool) is actually more useful because perhaps there's a part of the problem texts that we want to focus on at a particular time, we can look at the Apps that really relate to that and it's all kind of laid out and organized and so I think that's helpful and more useful than just the original text. - P2

Parsing recommended ideas into different categories offers more efficient navigation to find relevant ideas than just clustering all relevant ideas by semantic similarity as the sole measure.

Design Insight #2: Prioritize schema dimensions that highlights ideas' unique features. The participants have reported using a problem-solution division and looking for various solutions that tackle similar problems. The schema helped them find relevant information in a multi-faceted way so they could view a more refined list of ideas.

> I thought that was helpful to break up the task into all the different subcomponents, so that we can get something for each thing. Breaking it up really just diversifies the number of things we're looking at. - P2

However, in contrast to the previous research, the participants have also reported that individual examples may prioritize metadata on some dimensions over the others. Showing representatives of the unique features instead of more generic information can help users further narrow down results. Instead of representing the ideas with a universal structure where each metadata is treated equally and users should go over every dimension, the ability to highlight more unique features of an idea can help users further narrow down relevant ideas.

Design Insight #3: Diversify metadata on different schema dimensions for the same idea. The participants mentioned that the metadata was not useful when being too specific or too generic. The metadata might be too long to read or too specific to associate with other ideas, so it failed to serve as an exploratory cue.

> Sometimes it did feel like it was a little too long and it was just like running off of nothing. Like it was relevant, but it was still like it's just kind of just putting it there. - P4

On the other hand, generic metadata (i.e., "everyone") failed to narrow down relevant ideas for people to explore. Different dimensions of metadata might also overlap with each other for being too generic. Therefore, even though sometimes the workers answered the questions correctly in the crowdsourcing tasks, the collected metadata might not provide much value for inspiration seekers to take advantage of these metadata.

> Just like what kind of problem they're all trying to solve is like the unifying idea between all of like so something has to do with productivity, then I wanted to shake productivity. - P5

When collecting metadata from the crowdworkers, researchers should take procedures to ensure that the metadata for the same idea on different dimensions are specific enough, so they provide exploratory cues and diverse information.

References

Chan, J.; Chang, J. C.; Hope, T.; Shahaf, D.; and Kittur, A. 2018. Solvent: A mixed initiative system for finding analogies between research papers. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW): 1–21.

Chang, J. C.; Hahn, N.; and Kittur, A. 2020. Mesh: Scaffolding Comparison Tables for Online Decision Making. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 391–405.

Gentner, D.; and Markman, A. B. 1997. Structure mapping in analogy and similarity. *American psychologist* 52(1): 45.

Hope, T.; Chan, J.; Kittur, A.; and Shahaf, D. 2017. Accelerating Innovation Through Analogy Mining. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, 235–243. New York, NY, USA: Association for Computing Machinery. ISBN 9781450348874. doi:10.1145/3097983.3098038. URL https://doi.org/10.1145/3097983.3098038.

Hope, T.; Tamari, R.; Kang, H.; Hershcovich, D.; Chan, J.; Kittur, A.; and Shahaf, D. 2021. Scaling Creative Inspiration with Fine-Grained Functional Facets of Product Ideas. *arXiv preprint arXiv:2102.09761* .

Lopez-Vega, H.; Tell, F.; and Vanhaverbeke, W. 2016. Where and how to search? Search paths in open innovation. *Research policy* 45(1): 125–136.

Xu, X.; Fan, J.; and Dow, S. 2020. Schema and Metadata Guide the Collective Generation of Relevant and Diverse Work. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 8, 178–182.