Leveraging General Knowledge Graphs in Crowd-powered Innovation

Maximilian Mackeprang
Dr. Abderrahmane Khiat
Prof. Dr. Claudia Müller-Birn
maximilian.mackeprang@fu-berlin.de
abderrahmane.khiat@fu-berlin.de
clmb@inf.fu-berlin.de
Human-Centered Computing
Freie Universität Berlin
Berlin, Germany

ABSTRACT

Despite the significant benefits of crowd ideation platforms, these introduced two main challenges: (1) many ideas generated are basic and repetitive (2) the high number of ideas generated makes it practically impossible for ideation experts to review ideas one by one in order to select novel and useful ones. A key feature to overcome these issues resides in understanding the ideas. General Knowledge Graphs describe the meaning of domain-independent terms in an computationally understandable way and therefore represent a promising solution in obtaining such meaning. In this paper, we describe our research in understanding ideas and preliminary findings in application for crowd-powered innovation.

CHI'19, May 2019, Glasgow, United Kingdom

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

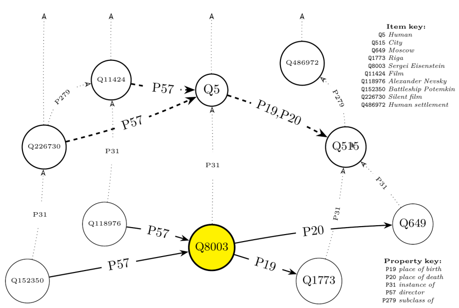

Figure 1: Example of knowledge organization in Wikidata (taken from [2])

CCS CONCEPTS

• **Information systems** → *Crowdsourcing*; • **Human-centered computing** → *Collaborative and social computing*; • **Computing methodologies** → *Semantic networks.*

KEYWORDS

Creativity, Crowd Ideation, Collaborative Ideation, Knowledge Graphs

ACM Reference Format:

Maximilian Mackeprang, Dr. Abderrahmane Khiat, and Prof. Dr. Claudia Müller-Birn. 2019. Leveraging General Knowledge Graphs in Crowd-powered Innovation. In *Proceedings of ACM CHI Conference (CHI'19).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

MOTIVATION

Innovation is one of the key driving forces of knowledge-based societies. In order to get heterogeneous ideas from diverse backgrounds, more and more companies turn to large scale open ideation platforms as a way to obtain creative ideas. On these platforms a crowd of ideators is asynchronously generating large amounts of ideas, that subsequently are filtered by experts.

However, two main challenges are introduced by these platforms: First, the submitted ideas are often simple, mundane and repetitive. Second, the amount of ideas makes it economically undesirable to evaluate them all manually [1]. To tackle the first problem, existing approaches guide the crowd during the idea generation phase by employing inspiration (e.g. showing similar or diverse ideas) as a creativity-enhancing intervention to enhance the quality of generated ideas. On the other hand, to support experts in selecting promising ideas, existing approaches employ statistical information extraction techniques such as word embedding based similarity computation, topic modelling and clustering. These statistical approaches are limited in two ways: Statistical evaluation of short text snippets is complex (and result quality varies depending on context and domain) and only accounts for surface features between texts, prohibiting a deeper understanding of underlying concepts used in the ideas.

In order to tackle these problems, we believe knowledge graphs offer a promising solution to (1) extract structured information about the content of ideas and (2) aid ideation experts during the further processing of result ideas by providing multi-faceted visualization of the ideation outcome.

BACKGROUND: KNOWLEDGE GRAPHS

In order to communicate our vision of an idea-based knowledge graph, we subsequently explain basic concepts used in the approach: A *knowledge graph* organizes various topically different real-world entities (concepts) with their relations in a graph. It provides a schema that aggregates these

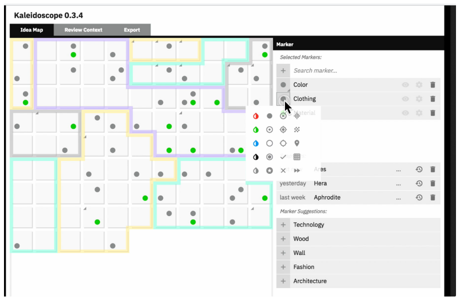

Figure 2: The Kaleidoscope user interface: Marking patterns in an idea space (taken from [7])

real-world entities into classes (abstract concepts) that also have relationships to each other [9]. In other words, using a knowledge graph provides computational meaning to concepts. A *general knowledge graph* encodes domain independent facts about concepts and entities (For example: "A dog is an animal") in large databases. Historically general knowledge graphs were built by domain experts, for example the *cyc* knowledge base [5], or *wordnet* [8]. In contrast, recent approaches moved from expert-curated systems to an open, collaborative knowledge graph creation approach: Most prominently, the *Wikidata* project, which was started in 2012 as a way to provide structured data for Wikimedia projects to overcome the data inconsistencies of Wikipedia's language versions provides structured interlinked data about entities [10]. An example of the graph format of Wikidata is shown in Figure 1. A hybrid approach to data creation is the *Dbpedia* project, that creates a general knowledge graph by automatically parsing information provided by English Wikipedia pages [4]. All of these examples provide different data about concepts that could potentially help making ideas computationally understandable.

APPROACH

Understanding ideas is a key feature of both delivering better inspiration (i.e. showing similar or diverse ideas) to the crowd during ideation, and better classification and assessment of ideas (i.e. clustering, topic modeling) in subsequent refinement steps. In order to obtain the meaning about the idea content, we employ an NLP based concept annotation mechanism, that is supplemented by an interactive error correction step [3], in order to store the correct meaning of the concepts used in ideas. More precisely, our approach uses part-of-speech tagging to extract terms in the ideas and link them to concepts in Dbpedia and Wikidata. We evaluated the impact of an additional task (i.e. validating concepts) in a brainstorming process [6]. During this task, users disambiguate terms in their ideas to link them to concepts stored in a general knowledge graph. The results obtained show no significant impact on ideation outcome.

By linking terms in ideas to knowledge graph concepts, the approach enables this meaning to be used as an analysis approach. As a first exploration of a potential application in sense-making of ideation outcomes, we developed an interactive visualisation prototype [7]. In this artifact, users can import ideas and visualize them in a grid. The users then can interactively select markers for concepts used in the ideas to visualize patterns within the ideation outcome (a prototype screenshot is shown in figure 2).

VISION

In order to enhance and augment crowd-powered large scale ideation, we envision an idea-based knowledge graph, leveraging semantic meaning about ideas to model multi-dimensional relations between ideas. Furthermore, by linking not only relationships between different ideas, but also

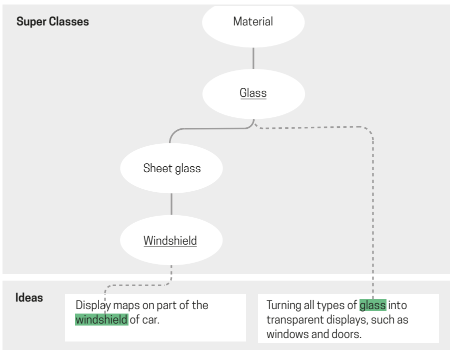

Figure 3: Example of how terms in ideas are linked to concepts in a knowledge graph.

relationships between different versions of the same idea, we envision a computationally accessible *Idea Lifecycle*, capturing iterations, refinement and different branches of ideas created during an innovation process.

When applied in a crowd ideation scenario, this idea-based knowledge graph could provide information about categories explored by the ideators and lead ideators to not well explored categories, help select tailored inspirations depending on information gathered for each ideator and the overall problem context, and aid in interactive visualization of the ideation outcome.

One promising use case for this idea-based knowledge graph is the computational analysis of inspiring example ideas, shown to crowd workers during ideation: In preliminary studies, we could show that inspirations are integrated into the ideators state of mind either directly (re-using the same term), indirectly (using a super-class of a term), or via the transfer of the *use* of a term. Figure 3 shows an example of an indirect link between terms in two ideas, based on concept super-classes extracted from Wikidata.

AT THE WORKSHOP

Combining the research areas of general knowledge graphs and large scale crowd-based ideation show great potential to help with the pressing issues in open innovation. In our contribution, we will present the basics of knowledge graphs using Wikidata as an example. By discussing our approach with a mix of different researchers and practitioners, we hope to advance this research area by focusing on the following questions:

How can we use knowledge graph data for adaptive creativity enhancing inspirations? In order to enhance crowd-powered ideation processes, a key aspect is the provision of creativity enhancing inspirations. Leveraging knowledge graphs by linking them with insights from cognitive science would enable a structured approach to the design and way of delivery for inspirations.

Which other ways are there to leverage knowledge graphs in crowd-powered ideation? In this paper, we have shown how to leverage knowledge graphs to a) Build an interactive sense-making tool based on a map metaphor and b) Analyze the integration of exemplars into subsequent ideas by means of super-classes extracted from Wikidata. We believe that general knowledge graphs hold many more potential applications for crowd-powered ideation. We envision an open discussion about the use of knowledge graphs in crowd-powered ideation systems in all related areas. By discussing speculative application scenarios we can map the design space and focus efforts on promising solutions.

ACKNOWLEDGMENTS

This work is supported by the German Federal Ministry of Education and Research, grant 03IO1617 ("Ideas to Market").

REFERENCES

[1] Joel Chan, Steven Dang, and Steven P Dow. 2016. Comparing different sensemaking approaches for large-scale ideation. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 2717–2728.

[2] Jakob Höper and Claudia Müller-Birn. 2018. Assisting in semantic enrichment of scholarly resources by connecting neonion and Wikidata. (2018).

[3] Abderrahmane Khiat, Maximilian Mackeprang, and Claudia Müller-Birn. 2017. Semantic Annotation for Enhancing Collaborative Ideation. In *Proceedings of the 13th International Conference on Semantic Systems (Semantics2017)*. ACM, New York, NY, USA, 173–176. https://doi.org/10.1145/3132218.3132235

[4] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. 2015. DBpedia–a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web* 6, 2 (2015), 167–195.

[5] Douglas B Lenat, Mayank Prakash, and Mary Shepherd. 1985. CYC: Using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks. *AI magazine* 6, 4 (1985), 65.

[6] Maximilian Mackeprang, Abderrahmane Khiat, and Claudia Müller-Birn. 2018. Concept Validation during Collaborative Ideation and Its Effect on Ideation Outcome. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, LBW033.

[7] Maximilian Mackeprang, Johann Strama, Gerold Schneider, Philipp Kuhnz, Jesse Josua Benjamin, and Claudia Müller-Birn. 2018. Kaleidoscope: An RDF-based Exploratory Data Analysis Tool for Ideation Outcomes. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*. ACM, 75–77.

[8] George A Miller. 1995. WordNet: a lexical database for English. *Commun. ACM* 38, 11 (1995), 39–41.

[9] Heiko Paulheim. 2017. Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic web* 8, 3 (2017), 489–508.

[10] Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 57, 10 (2014), 78–85.