

Wahl, Julian; Füller, Johann; Hutter, Katja

**Article**

## AI-Assisted Searching Through Crowdsourced Solution Space

Marketing Review St.Gallen

**Provided in Cooperation with:**

Universität St. Gallen, Institut für Marketing und Customer Insight

*Suggested Citation:* Wahl, Julian; Füller, Johann; Hutter, Katja (2022) : AI-Assisted Searching Through Crowdsourced Solution Space, Marketing Review St.Gallen, ISSN 1865-7516, Thexis Verlag, St.Gallen, Vol. 39, Iss. 6, pp. 30-38

This Version is available at:

<https://hdl.handle.net/10419/279710>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Marketing Review St.Gallen

## Open Innovation & Crowdsourcing



### Schwerpunkt

Mensch oder Maschine –  
Wer behält die Vorherrschaft  
bei der Kreativität?

«Golden Crowd» –  
Engaging the Users That Deliver  
on Your Crowdsourcing Goals

Beware of Sore Losers!  
Crowdsourcing Might Have Backfiring  
Effects Among Participants

AI-Assisted Searching Through  
Crowdsourced Solution Space

Human-Machine Creativity –  
How AI Can Influence Human  
Creativity in Open Innovation

### Spektrum

B-to-B-Marketing ist  
Hochleistungsmarketing

Rethinking Value Creation  
in Brick-and-Mortar Retailing

Believe the Hype?  
Herausforderungen und  
Herangehensweisen  
an das «Metaverse»



# AI-Assisted Searching

Through Crowdsourced  
Solution Spaces

In crowdsourcing contests, managers have to process hundreds of potential solutions on different topics and from different perspectives. This is resource-intensive, and the cognitive limits of humans are quickly exceeded. Applying contemporary AI-based language models is a promising way to help structure and explore crowdsourcing solutions.

Julian Wahl, M.Sc., Prof. Dr. Johann Füller, Prof. Dr. Katja Hutter

Hardly anyone engaged in the field of open innovation would deny that it is beneficial to have access to many different ideas in the front-end of new product development processes. But innovation managers that have to deal with the large number of ideas shared in crowdsourcing contests (Afuah & Tucci, 2012; Terwiesch & Xu, 2008) and extricate the potential of the shared suggestions may not always be that enthusiastic. Reading hundreds or even thousands of ideas is likely to lead to high cognitive load levels, negatively affecting decision quality (Blohm et al., 2016; Sweller & Chandler, 1994). Insights into the structure of the crowdsourced solution space (Jeppesen & Lakhani, 2010; von Hippel & von Krogh, 2016) may help to absorb essential information and better solve the broadcasted innovation problem. However, innovation managers tend to lose sight of essential and promising solutions and overlook important patterns in the gathered content. Thus, new approaches are needed to capture the full range of proposed solutions and overcome limited human processing capacities.

Integrating artificial intelligence (AI) into new product development allows for efficient processing of large-scale solution-related information and rethinking entire innovation processes (Füller et al., 2022; Haefner et al., 2021). In this context, applying AI-based language models presents a promising way to analyze large amounts of text. With these models, texts can be transformed into numeric vectors and allocated at specific locations in a high-dimensional space depending on their semantic content. Recently, pre-trained contextual language models that learn the semantic representations through transformer network architectures have revolutionized the field (Wolf et al., 2020). For example, sentence-BERT (SBERT) models (Reimers & Gurevych, 2019) based on Google's popular BERT model (Devlin et al., 2019) are tuned for automatically mapping entire sentences or paragraphs into a semantic space.

While there have been initial efforts to enrich the wisdom of the crowd with AI-based language models (Rhyn & Blohm, 2017), little research has explored how pre-trained contextual models for text representation can be implemented to gain insights into the structure of crowdsourced content and support human decision-making. Thus, in collaboration with Hyvecrowd – a leading German crowdsourcing platform – we explored how SBERT and complementary algorithms can be applied to facilitate AI-assisted searches through crowdsourced solution spaces. Using data from a crowdsourcing contest on future mobility and energy services, we show how AI-assisted searching can support managers to automatically identify clusters of similar ideas giving insights into their popularity, uncover latent topics and important interrelations, or directly spot ideas that meet a specific need or sub-problem. Thereby, it reduces the time and cognitive effort required to analyze the vast amount of submitted ideas.



**Julian Wahl, M.Sc.**

Research Assistant and PhD Candidate,  
University of Innsbruck, Austria  
[julian.wahl@uibk.ac.at](mailto:julian.wahl@uibk.ac.at)

**Prof. Dr. Johann Füller**

Professor of Innovation and Entrepreneurship,  
University of Innsbruck, Austria and  
CEO of HYVE – the innovation company,  
Munich, Germany  
[johann.fueller@uibk.ac.at](mailto:johann.fueller@uibk.ac.at)

**Prof. Dr. Katja Hutter**

Professor of Innovation and Entrepreneurship,  
University of Innsbruck, Austria  
[katja.hutter@uibk.ac.at](mailto:katja.hutter@uibk.ac.at)

## Literature Background

### Crowdsourced Solution Spaces

Compared to situations where organizations search for new solutions by themselves, crowdsourcing allows for richer insights, as the various contributors have information on different parts of the entire solution space (Jeppesen & Lakhani, 2010; von Hippel & von Krogh, 2016). Problem-solving literature suggests that the generation of multiple ideas can enable new ways of thinking (Boden, 2004; Ulrich, 1988). Therefore, decision-makers should not limit themselves to analyzing single search paths but rather consider a broad collection of possible solutions to find the right one (Simon, 1996). However, when organizations review crowdsourced idea content, they often focus on filtering out a specific number of winning contributions (Merz, 2018). As a result, a myriad of perspectives – representing the crowd's collective intelligence – is not considered in further decision-making. Yet, insights into that knowledge structure can be helpful to learn from the shared input and explore promising solutions.

Humans quickly reach their cognitive limits when processing many different ideas with multiple interacting components. If too many ideas are delivered at a particular time, the incoming information cannot be sufficiently organized and recognized according to its thematic content (Hiltz & Turoff, 1985; Sweller & Chandler, 1994). Thus, innovation managers engaged in crowdsourcing contests often struggle to develop a clear picture of the gathered knowledge. Although human experts can draw on a high level of domain-specific knowledge and experience, they

## Management Summary

This study explores how AI-based language models can complement the search through large numbers of solutions shared in a crowdsourcing contest. It outlines four search applications that automatically aggregate semantically similar ideas into clusters to explore characteristic agglomerations in the solution space, unveil underlying themes and interrelations to build a holistic understanding of the solution space, and identify solutions that address a pressing need or sub-problem within a few seconds without any noteworthy costs or cognitive effort. Innovation managers are encouraged to implement AI-assisted searching to improve the exploration of solution spaces in crowdsourcing contests and beyond.

need to find ways to overcome their limited processing capacities and develop a holistic understanding of the vast space of possible solutions to identify valuable crowdsourced contributions.

## AI-based Idea Representations

Recently, insights gained from analyzing large amounts of text with the help of language models for semantic text representation have gained traction in innovation and crowdsourcing research (Antons et al., 2020; Rhyn & Blohm, 2017). Since ideas shared in contests are described in text, processing ideas with language models powered by deep learning algorithms is predestined to uncover semantic similarities.

In the last decade, the application of deep learning algorithms to generate AI-based text representations has pushed the field of natural language understanding forward. Popular models like Word2Vec (Mikolov et al., 2013) or GloVe (Pennington et al., 2014) are learning semantic embeddings for words depending on how frequently they appear close to each other in a text corpus. Consequently, semantically similar words obtain similar vector values in the embedding space. Some models map entire sentences or short documents in an semantic space (Kusner et al., 2015; Le and Mikolov, 2014). For example, in the models the sentence “We published an article on open innovation in the Marketing Review St. Gallen” will have a similar semantic vector representation as “We wrote a paper on crowdsourcing practices for the Harvard Business Review.” However, distributional models for text representation like Word2vec or GloVe ignore the different meanings of words in different contexts. For example, a processed word like “paper” may encode different information

depending on whether it is used by an academic researcher or by an industry expert on packaging solutions.

Unlike distributional word representation models in which each word of a vocabulary corresponds to a single word embedding, contextual models consider the environment of each word occurrence. Thus, the value of the word vector varies depending on the sentence or paragraph in which the word is used. The capturing of the context is enabled by powerful transformer architectures, which can generate universal representations of an unprecedented amount of text data (Wolf et al., 2020). Contextual models such as BERT (Devlin et al., 2019) or GPT-3 (Brown et al., 2020) are pre-trained on large language datasets scraped from web corpora, such as Wikipedia, Gigaword, Google News, and many more. Pre-trained models using the SBERT framework (Reimers & Gurevych, 2019) deliver state-of-the-art embeddings for entire sentences or short documents.

The benefits of applying modern AI-based language models to automatically create similarity mappings of numerous crowdsourced ideas are apparent, as they can help innovation managers structure solution-related content automatically and extend their information processing capacity. However, the representation of potential solutions based on their semantic content is of limited value without search applications that support them in exploring the emerging solution space.

## Research Context and Data

In collaboration with the leading German crowdsourcing platform Hyvecrowd, we experimented on how AI-based representations of hundreds of crowdsourced solutions implemented in practical applications can enrich search practice. For this purpose, we used data from an innovation contest, in which the car manufacturer BMW reached out to the crowd to identify mobility and energy services that will potentially be established by 2030. More specifically, BMW wanted to know which services are changing the mobility ecosystem, which startups satisfy emerging customer needs, and which developments are in the early stages and have the potential to become established within the next ten years. With the chance to win prizes of € 3,000 in total, 151 participants generated 289 potential solutions.

To better process the heterogeneous perspectives shared by the crowd and navigate through the large amount of ideas, we applied state-of-the-art SBERT models from the Python package Sentence Transformers (Reimers, 2021) to automatically generate a similarity allocation of the 289 idea vectors. In several experiments, we explored how this allocation can support decision-makers in organizing the wisdom of the crowd more efficiently and effectively. Together with practical experts from



the team of Hyvecrowd and backed by academic concepts on solution search, we came up with four different approaches to navigate through crowdsourced knowledge that can be used sequentially: 1) exploring density structures; 2) determining topics and interrelations; 3) combining density structures and topics; 4) matching ideas with specific needs.

## Search Applications

### Exploring Density Structures

We assume that in exploration processes insights into the structure of the solutions space (Newell & Simon, 1972) facilitated by an automated allocation of ideas according to their semantic meaning can significantly support the search for solutions. Structuring the crowdsourced solution space by clustering the embedded idea vectors allows for a first understanding of

the characteristics of the shared set of solutions and facilitates processing the numerous possible solutions. Innovation managers often start their search by simply exploring exciting ideas. Using density-based clustering, managers can learn whether a proposed solution is part of a popular or unique neighborhood. For example, dense agglomerations might create an unbiased structure indicative of ideas addressing a relevant problem or need (Kornish & Ulrich, 2011).

In a hierarchical Ward clustering analysis (Scikit-learn, 2022), we uncovered 176 cluster groups that range from one to six solutions. Another clustering model with higher similarity threshold levels yielded 76 groups and revealed additional insights at broader levels of the crowdsourced solution space. Thus, two solution landscape structures on different levels of granularity emerged (figure 1). While the actual idea vector was generated based on a longer text, the subtitles validate the face validity of the generated similarity structure. For

Figure 1: Section of Clustered Solution Spaces on Two Different Granularity Levels

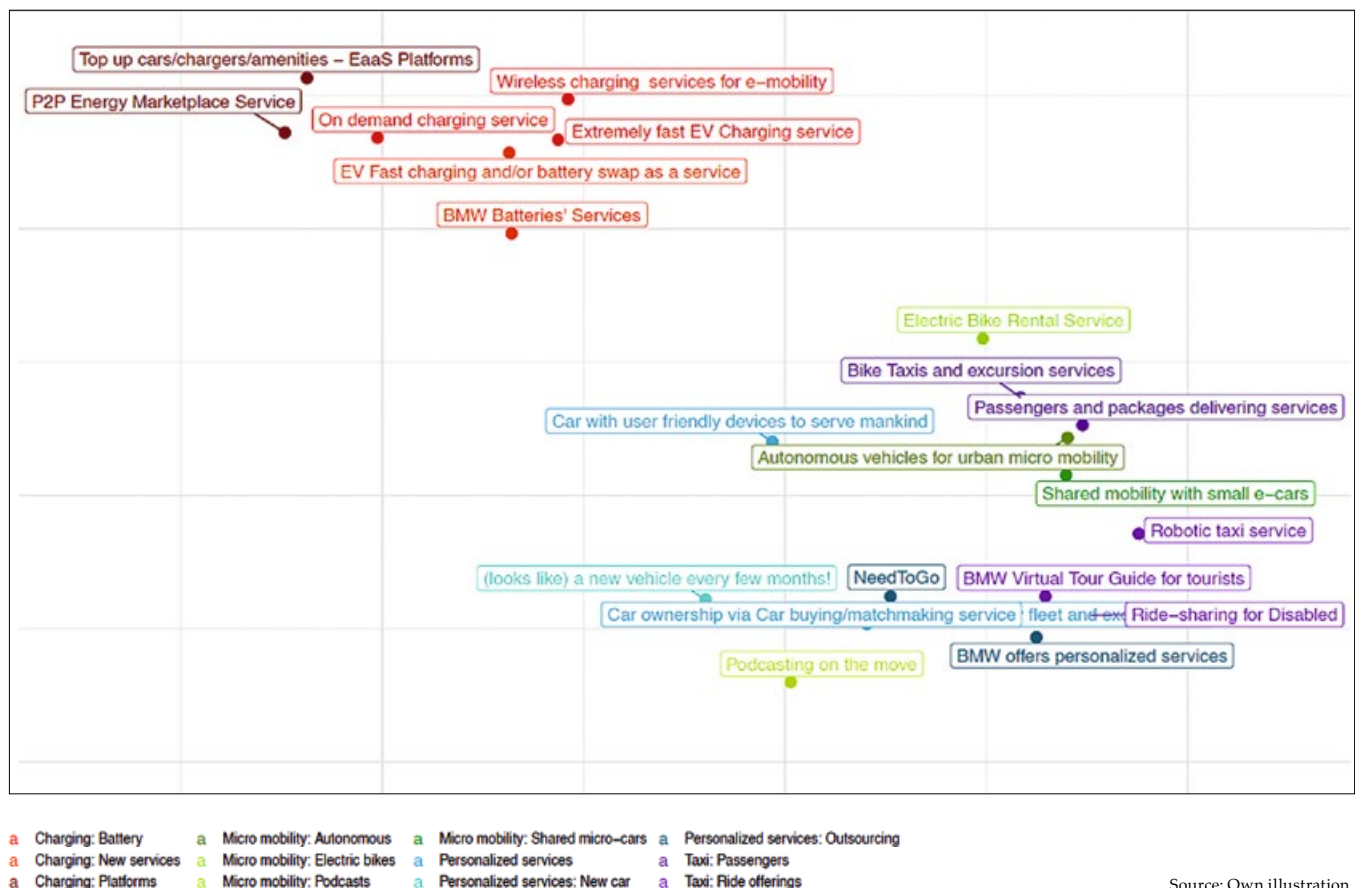
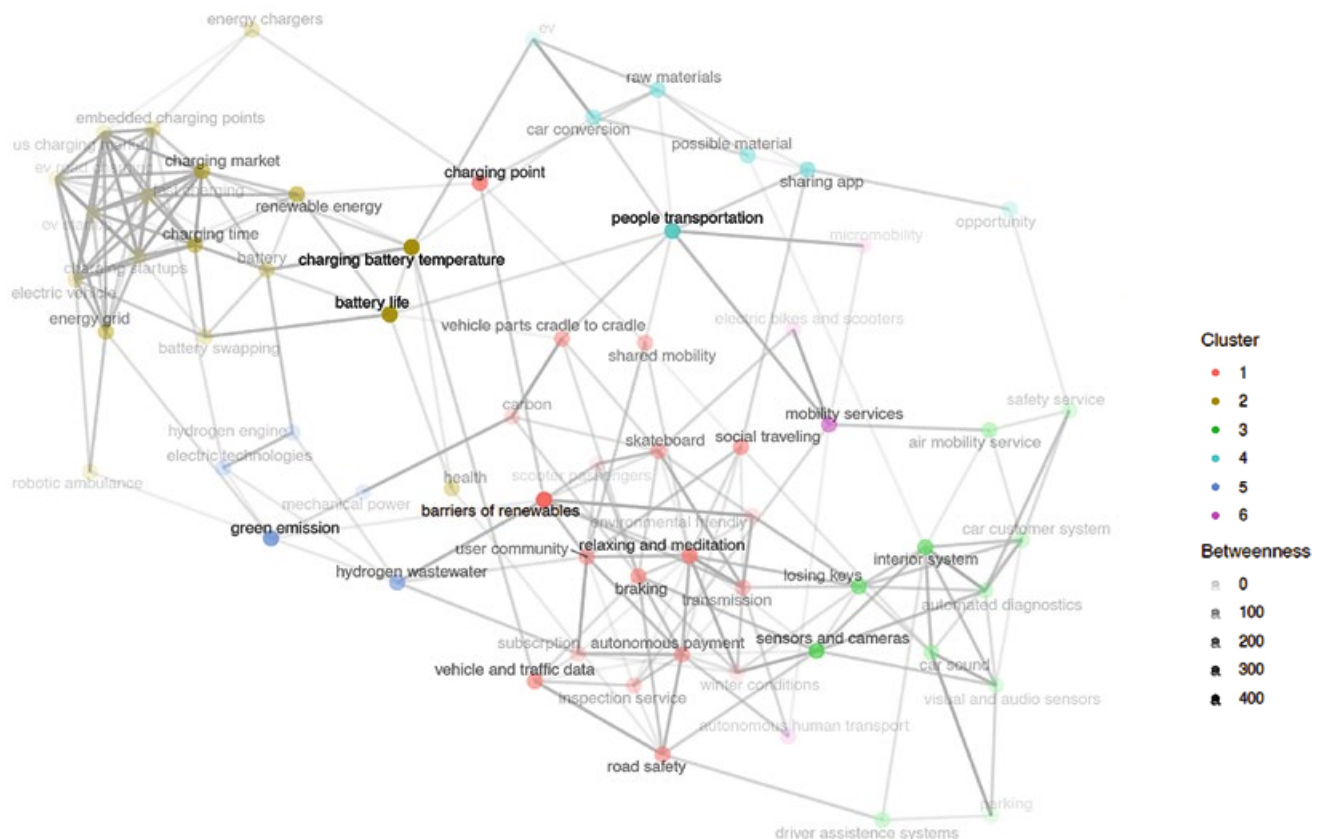


Figure 2: Topic Network



The network uses the correlations between the topic loadings of the crowdsourced service descriptions and only displays correlations above a threshold value of the 90th percentile. Source: Own illustration.

example, the red cluster in figure 1 involves solutions dealing with electric vehicle (EV) charging. This cluster can be divided into three further groups covering batteries, new charging services, and platforms. To obtain the representation, the 768-dimensional SBERT vectors of each service description were reduced to two dimensions using the UMAP technique (McInnes et al., 2018). This allows for more nuanced insights into the semantic density structure of the solution gathered in the crowdsourcing contest.

## Determining Topics and Interrelations

Besides semantic clusters, innovation managers engaged in crowdsourcing contests may also be interested in common themes independent of density structures in the crowdsourced solution space. Topic modeling based on SBERT embeddings allows the identification of interrelations between ideas that

share similar topics but might approach them from different perspectives (Bianchi et al., 2021). In the cluster exploration outlined above, the more latent interrelations might remain undetected. We ran a contextual topic model analysis over all 232 crowdsourced ideas using the Python package Contextualized Topic Models (Bianchi et al., 2020). For the analyzed data, a model with 70 topics turned out as most effective for revealing coherent, diverse, and fine-grained regularities among the shared mobility and energy service descriptions. Insights into the regularities across contributed solutions and their components may further transform the overall understanding of the innovation problem. We manually labeled the solutions by focusing on the topic words with the highest scores and removed six noisy topics with incoherent and meaningless data.

In the network shown in figure 2, the topics are linked based on their common occurrence in the crowdsourced service descriptions. The different colors of the nodes in the network represent

different topic communities, the different transparencies of the nodes indicate their importance, with the most opaque ones highlighting the connector nodes in the network. The thickness of the connecting lines corresponds to the strength of the correlation between the topics. For example, in the top left corner of the figure, the light brown topics deal with charging and electricity services around the car of the future and indicate exceptionally high correlations. This topic community further involves the two important topics of charging battery temperature and battery life that connect it to other communities. Innovation managers can learn from the unveiled patterns and understand the solution space, including underlying sub-problems, important interrelations, and potential components of solutions. In a further step, nuanced and holistic representations of possible ways to solve the crowdsourced problem may become more transparent. For example, topic combinations may serve as a template that stimulates the search for new solutions beyond the ones created in the crowdsourcing contest.

## Combining Density Structures and Topics

When one has information about the semantic mapping, density structures, and underlying themes of the crowdsourced solution space, combining this information into a single search application makes sense. Figure 3 illustrates an interactive visualization of the crowdsourced solution space in three-dimensional space. This allows to roam through all 289 service

descriptions and gain information on their content, popularity in the contest, i.e., likes and comments, and the density in the semantic neighborhood, i.e., the number of contributions in the fine and broad clusters. While dense clusters indicate popular solutions in the neighborhood, smaller clusters may reveal unique posts that indicate weak signals, e.g., in the case of a high number of views, comments, or likes. Using this information, decision-makers can focus on selected areas of the solution space or read through these potential solutions in more detail.

Furthermore, by connecting the most characteristic topics of all service descriptions, one can filter for overarching topics of interest and focus one's search on specific parts of the crowdsourced solution space, starting from a topic of interest. For example, the service description "Autonomous vehicles for autonomous urban micro mobility" depicted in the lower graph of figure 3 strongly represents the topic of autonomous human transport. However, compared to the service description "Powerbank – mobile charging on the road" shown in the upper graph, it appears less attractive, with lower values in comments and likes, and a fine and broad cluster size.

## Matching Ideas with Specific Needs

Innovation managers are often aware of specific problems or are searching for ideas addressing specific issues. In this case, they may search through manually annotated idea tags

### Main Propositions

- 1 Managers require new approaches to overcome cognitive limitations in processing large-scale text information.
- 2 The automatic clustering of semantically similar solutions creates an unbiased structure to identify popular or unique contributions.
- 3 Topic modeling helps to better understand underlying components in crowdsourced solution spaces.
- 4 Instead of keywords or themes, customized need queries can be used to identify relevant solutions directly.
- 5 Though many tasks can be automated, innovation managers' domain knowledge and contextualization remain essential.

### Lessons Learned

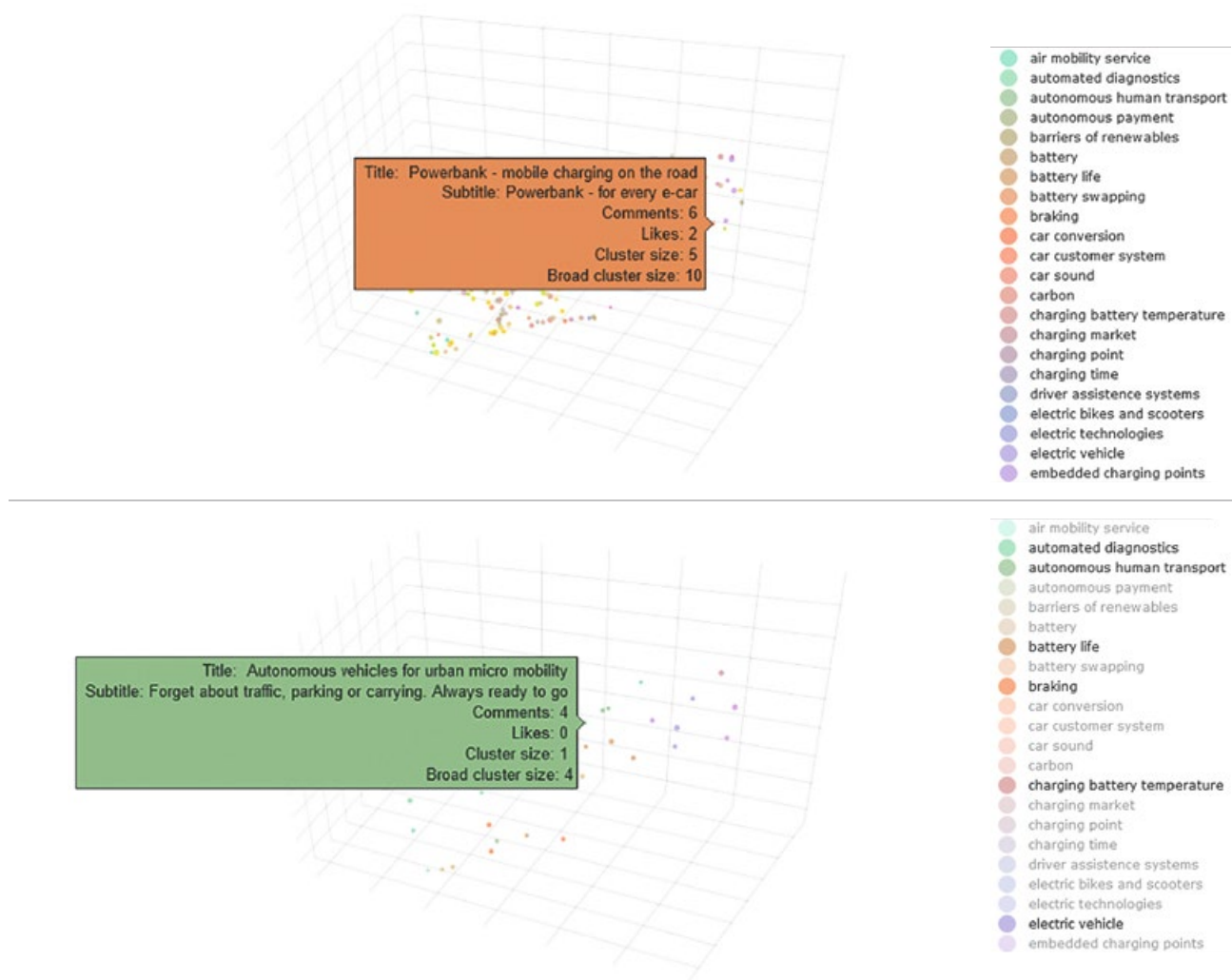
- 1 The different search applications facilitated by AI-based language models should be combined into scalable tools and integrated with idea management and crowdsourcing platforms.
- 2 Innovation managers are encouraged to let AI handle any tasks that involve the processing of large amounts of text to save resources for high-value tasks.
- 3 Before implementing AI-assisted searching, innovation managers should carefully validate the benefits of previous practices and clarify where experts' intuition and experience are required.
- 4 AI-savvy innovation managers should closely follow the rapid advancements in the automated analysis of text, images, and videos and constantly explore new options to enhance AI-assisted solution searches.



or filter for related keywords in the crowdsourced idea titles to limit the amount of text to be investigated. However, idea titles or annotated idea tags only capture a limited part of the semantic content. For example, keywords that are not part of a pre-defined search set cannot be considered, and annotated tags and themes are likely too broad to engage in nuanced fine-grained searches. Due to their ability to capture semantic meaning and the context of language, SBERT models are promising tools to search for solutions that address pre-defined needs independent of certain keyword expressions, tags or themes by comparing the semantic distances between text vectors describing needs and solutions.

In the background information of the crowdsourcing contest, the hosting car manufacturer explicitly mentioned this issue. In the original contest, participants annotated their contributed ideas with pre-defined terms but could only select between the extremely broad trend themes of “mobility services” and “energy services”. In another step, Hyvecrowd and BMW experts invested considerable resources to categorize the contributed ideas into a few somewhat more specific themes, e.g., “maintenance service”, “charging service” or “human-machine interaction”. As these annotated themes still cover quite broad need areas, such a procedure is useless in a deliberate search for specific needs. In the analyzed contest briefing, the hosts men-

Figure 3: Exemplified Exploration of 3D Crowdsourced Solution Spaces



Source: Own illustration.

tioned needs like “How can we integrate cars into overarching mobility concepts?” or “Which data-based services can support the driver, fleet manager or government?”. Table 1 displays the results from two exemplary searches using these two need queries to automatically identify the ten crowdsourced ideas with the least semantic distance to them. Though the calculations are based on the full-text descriptions, a first evaluation of the results’ titles and subtitles confirms the search approach’s applicability and content validity. Almost all top-ten closest solutions provide a valid answer to the stated queries. Only two of the 20 reported outputs (listed in *italics*) do not provide useful answers. For example, the idea of seeing a BMW as a home away from home is not a solution for integrated mobility concepts, and the application of new technologies for seamless car rental may only indirectly be related to data-based services.

## Discussion

Usually, managers have to read through every idea submission to explore new and relevant ideas, break down idea content into the underlying components and identify solutions that match their needs. Our research outlines four different AI-based approaches to navigate the solution-related input shared in crowdsourcing contests. The proofs-of-concept described in this article illustrate the potential of AI-based language models and related algorithms to facilitate new ways of organizing the wisdom of crowds more efficiently and effectively. Previous research has already outlined how AI transforms innovation practices and which affordances emerge from AI-based innovation management (Füller et al., 2022; Haefner et al., 2021). In AI-assisted searches, large crowdsourced solution spaces can be structured in just a few seconds without any noteworthy costs or cognitive efforts. While the exploration of dense agglomerations reveals unbiased structures that might indicate solutions that address a relevant problem or need, smaller clusters can help find unique contributions (Kornish & Ulrich, 2011). Topic modeling supports gaining nuanced insights into latent themes and interrelations across all crowdsourced contributions and facilitates a holistic understanding of possible ways to solve the problem presented to the crowd. In addition, managers can focus on specific needs and directly identify solutions to address them, and even combine the different search applications one after the other.

While the proposed AI-based approaches examined in this study can complement the skills of human decision-makers by automating the structuring of idea content, the automation of entire innovation processes remains a thing of the future. In the investigated applications, human skills were important in conceptualizing approaches that provide value in practice, selecting suitable language models and parameters, and drawing conclusions from the generated output. While we manually labeled the

**Table 1: Top-10 Closest Crowdsourced Solutions of Need–Solution Matching (original wording)**

<b>Need query: “How can we integrate cars into overarching mobility concepts?”</b>	
1.	Opinionated #SoMe interactions, Social Audio ... – Socialize while on the go! Building on apps and internet connectivity, head up displays, etc.
2.	Shared mobility with small e-cars – getting where you want with your kids or friends in no time has never been that easy!
3.	The future is van-tastic – electric vans are conquering villages and towns.
4.	Service for product – carpooling, car sharing or other services.
5.	Distribution of condition(s) sensitive medication – collect your critical medication while you charge up!
6.	Mobility flat – vehicle tariffs as required.
7.	First-mile and last-mile mobility that stows easy and fits and recharges in your vehicle.
8.	Localelectric (local + electric) mobility services – providing local mobility services via electric vehicles.
9.	<i>Lounge BMW’s Services – electric car with more space and comfort. The EVs will be designed to be a home away from home.</i>
10.	BMW and the all-in-one solution – how to efficiently combine all sustainable trends? Configure-to-order process is the future!
<b>Need query: “Which data-based services can support the driver, fleet manager or government?”</b>	
1.	Data management from connected mobility – with vehicles acquiring more data from their drivers, managing and using this data is crucial to offer new services.
2.	Value from V2I, V2X and vehicle sensor data – vehicular or fleet sensing of surroundings and V2I, V2X, V2V touchpoints/history = trove of situational/environmental big data
3.	Sharing carsharing cars with the police App-open – keyless opening tech is developed into an app environment of additional services.
4.	Full system diagnostic scanner AI-Transport – damage detection for various service scenarios and vehicle diagnostics.
5.	EV mileage optimization service – EV vehicle, trip and charging management and optimization service to serve and help customers.
6.	Remotely operated/controlled EV – we all love remote controls, what about remotely operated/controlled EV?
7.	Smart vehicles with intelligence on board – a customizable vehicle assistance system helps to avoid personal and property damage – real(-time) assistance and sustainability
8.	Vehicle repair that is itself mobile – comes to you / your driveway / parking spot.
9.	<i>Premium shared mobility solutions – using technology to enable seamless car rental.</i>
10.	Geo targeted ads service – the digital billboards on the road syncs up their ads to what is relevant to the moving customers on wheels.

Source: Own illustration.

clusters and topics, future research is encouraged to experiment with language models for text generation (Ouyang et al., 2022) to automatically assign titles to clusters and topics based on the semantic content of the aggregated solutions or topic words.

Furthermore, in-depth knowledge about the organizational context, the market situation, or competitive scenarios appears indispensable to find the best overall solutions, at least for now. However, the days when experts relied solely on their cognitive abilities to make sense of vast amounts of crowdsourced knowledge to develop a thorough understanding of a problem and find the best solutions seem to be over. Innovation managers are encouraged to implement AI-assisted search approaches to improve the exploration of large amounts of ideas in crowdsourcing contests and beyond. Thus, they can reduce the cognitive demand associated with perusing every idea and find similar structures and interrelations that might have remained undetected with traditional approaches. In doing so, they should constantly monitor the advancements in existing AI-based language models accessible via open-source libraries on Python and ensure capabilities to implement them into their processes, either through customized tools or whole systems embedded within crowdsourcing or idea management platforms. ■

## References

- Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of Management Review*, 37(3), 355–375. <https://doi.org/10.5465/amr.2010.0146>
- Antons, D., Grünwald, E., Cichy, P., & Salge, T. O. (2020). The application of text mining methods in innovation research: Current state, evolution patterns, and development priorities. *R&D Management*, 50(3), 329–351. <https://doi.org/10.1111/radm.12408>
- Bianchi, F., Terragni, S., & Hovy, D. (2020). Python package “contextualized topic models.” <https://contextualized-topic-models.readthedocs.io/en/latest/>
- Bianchi, F., Terragni, S., & Hovy, D. (2021). Pre-training is a hot topic: Contextualized document embeddings improve topic coherence. *Proceedings of the ACL-IJCNLP 2021*, Vol. 2, pp. 759–766. <https://doi.org/10.18653/v1/2021.acl-short.96>
- Blohm, I., Riedl, C., Füller, J., & Leimeister, J. M. (2016). Rate or trade? Identifying winning ideas in open idea sourcing. *Information Systems Research*, 27(1), 27–48. <https://doi.org/10.1287/isre.2015.0605>
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). Routledge.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). Language models are few-shot learners. <https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 4171–4186.
- Füller, J., Hutter, K., Wahl, J., Bilgram, V., & Tekic, Z. (2022). How AI revolutionizes innovation management – Perceptions and implementation preferences of AI-based innovators. *Technological Forecasting and Social Change*, 178. <https://doi.org/10.1016/j.techfore.2022.121598>
- Haefner, N., Wincnet, J., Parida, V., & Gassmann, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change*, 162. <https://doi.org/10.1016/j.techfore.2020.120392>
- Hiltz, S. R., & Turoff, M. (1985). Structuring computer-mediated communication systems to avoid information overload. *Communications of the ACM*, 28(7), 680–689. <https://dl.acm.org/doi/10.1145/3894.3895>
- Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization Science*, 21(5), 1016–1033. <https://doi.org/10.1287/orsc.1090.0491>
- Kornish, L. J., & Ulrich, K. T. (2011). Opportunity spaces in innovation: Empirical analysis of large samples of ideas. *Management Science*, 57(1), 107–128. <https://doi.org/10.1287/mnsc.1100.1247>
- Kusner, M. J., Sun, Y., Kolkin, N. I., & Weinberger, K. Q. (2015). From word embeddings to document distances. *Proceedings of the 32nd International Conference on Machine Learning*, PMLR 37, 957–966. <https://proceedings.mlr.press/v37/kusnerb15.html>
- Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *Proceedings of the 31st International Conference on Machine Learning*, PMLR 32(2), 1188–1196. <https://proceedings.mlr.press/v32/le14.html>
- McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. *Arxiv*. <http://arxiv.org/abs/1802.03426>
- Merz, A. B. (2018). Mechanisms to select ideas in crowdsourced innovation contests – a systematic literature review and research agenda. *Proceedings of the 26th European Conference on Information Systems*. [https://aisel.aisnet.org/ecis2018\\_rp/68/](https://aisel.aisnet.org/ecis2018_rp/68/)
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations, Workshop Track Proceedings*, pp. 1–12. <https://arxiv.org/abs/1301.3781>
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Sandhini, Z., Katarina, A., Alex, S., John, R., Jacob, S., Fraser, K., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *OpenAI*. [https://cdn.openai.com/papers/Training\\_language\\_models\\_to\\_follow\\_instructions\\_with\\_human\\_feedback.pdf](https://cdn.openai.com/papers/Training_language_models_to_follow_instructions_with_human_feedback.pdf)
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. *Conference on Empirical Methods in Natural Language Processing*, pp. 1532–1543. <https://doi.org/10.3115/v1/d14-1162>
- Reimers, N. (2021). Python package “SentenceTransformers.”. SBERT. <https://www.sbert.net/docs/>
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, EMNLP-IJCNLP*. <https://aclanthology.org/D19-1410/>
- Rhyn, M., & Blohm, I. (2017). Combining collective and artificial intelligence: Towards a design theory for decision support in crowdsourcing. *Proceedings of the 25th European Conference on Information Systems*, 2656–2666. [https://www.alexandria.unisg.ch/252288/1/JML\\_660.pdf](https://www.alexandria.unisg.ch/252288/1/JML_660.pdf)
- Scikit-learn. (2022). Hierarchical clustering. <https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering>
- Simon, H. A. (1996). *The Sciences of the artificial* (3rd ed.). MIT Press.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 185–233. [https://doi.org/10.1207/s1532690xci1203\\_1](https://doi.org/10.1207/s1532690xci1203_1)
- Terwiesch, C., & Xu, Y. (2008). Innovation contests, open innovation, and multiagent problem solving. *Management Science*, 54(9), 1529–1543. <https://doi.org/10.1287/mnsc.1080.0884>
- Ulrich, K. T. (1988). *Computation and pre-parametric design*. Technical Report 1043. MIT Artificial Intelligence Laboratory. <http://hdl.handle.net/1721.1/6845>
- von Hippel, E., & von Krogh, G. (2016). Identifying viable “need-solution pairs”: Problem solving without problem formulation. *Organization Science*, 27(1), 207–221. <https://doi.org/10.1287/orsc.2015.1023>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, J., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., & Rush, A. (2020). Transformers: State-of-the-art natural language processing. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 38–45. <https://doi.org/10.18653/v1/2020.emnlp-demos.6>