Research topic evolution: a comparative analysis of human and machine approaches

Siyi Xiao✉ and Daniel A. McAdams

Texas A&M University, USA

✉ cindyx@tamu.edu

ABSTRACT: Exploring patterns in large text corpus is essential for effective knowledge discovery in research domains. However, machine-driven methods often introduce noise and rely heavily on parameter thresholds. Human expertise is therefore essential for ensuring reliable outcomes. This study conducts a comparative analysis of a classification task performed by both human and computer algorithms. During the task, human experts are asked to categorize a list of abstracts based on their semantic contents, where computer algorithms perform computations, including network analysis and document embeddings, to group the abstracts. The findings show a significant level of disagreement between human and computer-generated clusters, indicating the need for further investigation into the factors influencing community categorization and incorporating more advanced techniques to improve the results.

KEYWORDS: design methodology, design methods, case study, network analysis

1. Introduction

Applying network analysis and natural language processing to a large set of corpora has been widely explored for its potential to identify patterns and forecast trends. One common method is through embedding techniques, which transform texts into numerical representations and group similar ones together. Several studies have applied such techniques to find research trends (Huang et al., 2022; Lo & Wang, 2023), classify texts (Mehta et al., 2021; Wang et al., 2016), and identify topics (Mu et al., 2022). In our related work, we also used word embedding techniques and network analysis to visualize research trends within the design methodology community. Specifically, keywords extracted from the corpus are transformed into word vectors, which are then used to compute semantic similarities between keywords and construct networks based on these similarities. The networks are then divided into several communities where keywords with similar semantic meanings are grouped, and for each community, the most representative keyword is selected as a community label, which is considered a topic. Topic trends can subsequently be formed by calculating the similarities between topics in consecutive years (Xiao & McAdams, 2025).

While our related study provides a generalized approach to exploring trends within the engineering design methodology field, several challenges remain. The existing method does not eliminate all the noise data, causing general words to be chosen as a topic. Additionally, the current community detection method lacks robustness since the technique is highly sensitive to human-defined threshold values (Xiao & McAdams, 2025).

To address these gaps, this paper extends the previous study's framework by introducing new techniques for community and topic detection. Furthermore, the communities generated by computers are compared to those created by humans. This advancement avoids problems caused by irrelevant words extracted. Furthermore, comparing human and computer-generated communities helps evaluate the validity of the proposed method and provides deeper insight into the strengths and limitations of automated techniques.

2. Literature review

Citation and co-citation analysis are two types of network analysis frequently used to uncover relationships and explore patterns within scholarly literature. Citation analysis determines the importance of a paper through the number of citations it receives; therefore, it is usually applied to find the most influential paper in a field. Co-citation analysis is used to study how often two publications co-occur in the same bibliography, which is useful for identifying the underlying themes within the literature. In addition to citation and co-citation analysis, co-word analysis is also commonly used to determine patterns in large datasets. Similar to co-citation analysis, co-word analysis is based on the co-occurrence frequency of keywords in a collection of texts. Co-word analysis assumes two keywords appearing together share a similar thematic relationship and groups these keywords. The ability of co-word analysis to forecast future trends is derived from the content, as authors typically discuss their research contribution and future research within the article. By extracting keywords from the content, co-word analysis provides a preview of research trends. In contrast, citation and co-citation analysis are unable to achieve this, as these citation-based methods do not consider the paper's content. Since citation and co-citation analysis lack the ability to forecast future trends, researchers often combine co-word and co-citation analysis to enrich the findings and predict future research directions (Donthu et al., 2021). However, co-word analysis has drawbacks, such as words having multiple meanings and some terms being too general to categorize into a specific cluster. Therefore, novel approaches that involve using word embeddings to capture semantic meaning and calculate the semantic similarity of words, rather than a co-word network that relies on the co-occurrence frequency of words, have been proposed (Huang et al., 2022). A popular word embedding model, Word2Vec, is a multi-layer neural network that transforms words into vector representations. These word vectors can further be used to determine key themes within literature.

Similar to how network analysis integrates word embedding models for topic extraction, topic modeling can also leverage word embeddings to uncover latent structures in a corpus. Topic modeling is a technique based on the co-occurrence of keywords, however, instead of measuring the actual co-occurrence frequency of the words, the algorithm measures the probability of the co-occurrence (Chen et al., 2021), and groups extracted words according to a probability distribution (Leydesdorff & Nerghes, 2017). Latent Dirichlet Allocation (LDA) is one of the most well-known topic modeling techniques. The LDA model assumes each document contains a mix of underlying topics, which are determined by the *"probability distributions over words"* within the document (Vayansky & Kumar, 2020, p.3). Since traditional topic modeling algorithms group frequently co-occurring words together, the meaning of the words is not considered. In addition, LDA has some drawbacks, such as inconsistent results and ineffective assumptions. Therefore, many studies have proposed other topic models that derive from LDA or combine other statistical models with LDA to achieve better results (Vayansky & Kumar, 2020). Incorporating the word embedding technique with topic models is one of the methods. For example, Zhang et al. (2023) proposed an LDA2vec model where the LDA model was used to identify the topics first, and the topics were then transformed into vectors through the Word2Vec model (Zhang et al., 2023). With the multiplication of the topic vector and LDA topic probability value, a new topic vector is generated and used to compute the *"semantic similarity among topics at adjacency stages, and evolution paths that directly reflect the topic relationships are constructed"* (Zhang et al., 2023, p.1).

Topic modeling outputs a list of topics, where each topic consists of a mix of words. This nature of topic modeling reduces the interpretability of the results, as the output may not always be meaningful. Therefore, our study continues to use embedding techniques with network analysis as in our previous work. However, instead of transforming words into vectors, our new study transforms each abstract into a vector by using the Doc2Vec model. The Doc2Vec model, as an extension of the Word2Vec model, was developed by Le and Mikolov in 2014 to represent sentences and documents as vectors. In line with the Word2Vec model, the Doc2Vec algorithm also has two model architectures - the Distributed Memory Model of Paragraph Vectors (PV-DM) and the Distributed Bag of Words version of Paragraph Vector (PV-DBOW). The PV-DM model extends Word2Vec's CBOW model by incorporating a paragraph vector into the training process, which is then concatenated with word vectors to make predictions. The PV-DBOW model is analogous to Word2Vec's Skip-gram model but instead of predicting surrounding words given a word, the PV-DBOW model predicts words with a given paragraph vector (Le & Mikolov, 2014). To aid readers who are unfamiliar with the previously described techniques, Table 1 provides a summary of them. Word2Vec was employed in our previous study but demonstrated limited effectiveness. Therefore, Doc2Vec was chosen in this study to explore potential improvements.

Tabel 1. Summary of computational methods for topic analysis

Computational technique	Description	Pros/Cons
Word2Vec	An embedding method with two types of model architecture - CBOW and Skip-gram. Used to transform words into vector representations.	Capture semantic relationships
Doc2Vec	An extension of the Word2Vec model with two types of model architecture - PV-DM and PV-DBOW. Used to transform sentences and documents into vectors.	Preserve broader contextual information
LDA	A common topic modeling technique that assigns frequently co-occurring words into the same group. It can be used with embedding methods to enhance performance.	Require user to predefine a fixed number of topics

3. Research approach

This study adopts and extends approaches by Huang et al. (2022) (Huang et al., 2022) and the research framework builds on our related work (Xiao & McAdams, 2025), with modifications in the network construction and topic selection steps.

3.1. Network construction

The dataset analyzed here consists of the American Society of Mechanical Engineers (ASME) International Design Engineering and Technical Conference proceeding abstracts spanning five years, from 2018 to 2022, comprising a total of 268 papers. Text preprocessing techniques are applied to the dataset to eliminate unnecessary punctuation and words. Initially, the same dataset was used to train a Doc2Vec model, which generates document vectors for each abstract in the dataset. However, the resulting vectors are not diverse, making it a challenge to detect distinct communities. Therefore, more abstracts from previous years of the conference proceedings are included in the training process to avoid data overfitting. Each abstract is considered a single document, and once the documents are transformed into vectors, the cosine similarity is calculated between pairs of vectors to form a square similarity matrix. The square matrix is then used to construct networks where each entry (i,j) represents the cosine similarity between vectors i and j. To construct a network, a threshold is defined by the user such that an edge is added to the network only when the cosine similarity exceeds this threshold. Community detection algorithms are then applied to the resulting networks to divide them into groups, with each group consisting of closely related abstracts. Two programs are involved during the network construction and community detection process - NetworkX and VOSviewer (Xiao & McAdams, 2025). NetworkX is a Python package for network construction (Hagberg et al. 2008), and VOSviewer is used for visualization and community detection (van Eck & Waltman, 2014).

3.2. Topic identification and similarity calculations

Previous studies determine a single keyword with the highest Z-Score as a community label (Xiao & McAdams, 2025; Huang et al., 2022). Z-Score is an indication of node importance within a network, with a higher number representing higher importance. However, a single document (node) cannot adequately represent the entire document community, as each document is distinct and has its unique characteristics. Therefore, generalized topics are essential to better represent the entire community of documents. To obtain generalized topics, large language models (LLMs) are employed in our study. Specifically, ChatGPT determines the community labels based on the titles of the documents within each community, and these generated labels serve as the topics. The steps are as follows. First, a list of titles from a single community is input into ChatGPT with the prompt: "*Based on the list of titles I have provided, give me a label for this community of titles. The label should be less than 6 words.*" Secondly, the same procedures are repeated for the remaining communities within a single network until a list of LLM-generated topics is produced. Lastly, steps 1 and 2 are repeated for the remaining years of networks. However, to obtain the similarity between community labels in subsequent years and construct a Sankey diagram, it is still

necessary to calculate the Z-Score for each node and the similarity scores for all pairwise combinations of communities from consecutive years. Equations 1 and , adopted from Huang et al. (2022)'s work, demonstrate these two calculations.

$$z_i = \frac{N_M^i - \left(\sum_{(j \in M)} N_M^j \right)/M^o}{\sqrt{\sum_{(j \in M)} (N_M^j)^2/M^0 - \left(\sum_{(j \in M)} N_M^j/M^o \right)^2}} \tag{1}$$

$$Similarity(H(M_t), H(M_{t+1})) = \frac{\sum_{\substack{W_t \in H(M_t) \\ W_{t+1} \in H(M_{t+1})}} Z'_{W_t} Z'_{W_{t+1}} \cos(\nu_{W_t}, \nu_{W_{t+1}})}{\sum_{\substack{z_i \in Z'_t \\ z_j \in Z'_{t+1}}} z_i \cdot z_j} \tag{2}$$

In both equations, the letter z represents the Z-Score, and z' represents the max-min normalized Z-Score. N_M^i represents the total edge weights between the ith document and all other documents in community M, where M° represents the total number of documents in community M. Additionally, $H(M_t)$ represents the document set in community M at year t, and ν_{W_t} denotes the vector representation of document W_t (Huang et al., 2022). Once similarity scores are calculated for each pair of communities, a Sankey diagram can be created by forming paths between community pairs whose similarity scores exceed a user-defined threshold.

3.3. Human validation

Since categorization is generally a challenging task, this study also examines the algorithm's performance by investigating how documents are categorized by humans and comparing these results with those produced by computer algorithms. This human effort serves to validate the categorization process. Once the accuracy of the automated method is confirmed against this benchmark, human involvement may no longer be necessary for future categorizations.

The document categorization is done by five design methodology research field experts. Each person completed the categorization exercise independently. There were no instructions on how to group these documents because the study also aims to explore the diversity of researchers' decision-making processes. Each expert comes up with their number of clusters for the same dataset and their ways of grouping the documents. Table 2 summarizes each researcher's strategy to develop the document clusters. According to the table, each researcher performs the classification task using a different approach. Some researchers base their judgments on specific criteria, such as design goals or design stages, while others rely entirely on their intuition. The varying classification methods can lead to very divergent results. However, one of the goals of this paper is to investigate the extent to which decision-making differences impact the results and provide future study directions.

Tabel 2. Categorization strategies from each researcher

Researcher 1	Researcher 2	Researcher 3	Researcher 4	Researcher 5
Read the papers one by one, listing all the PDF file names in an Excel sheet and typing notes for each paper. After reading 15 papers, the researcher has a rough idea of how many clusters are in total. Then, the researcher groups all the papers at once, and several abstracts are assigned to more than one label.	Open 5 to 10 papers, one at a time, and read through them. If there are papers focused on the same theme, create a folder and move those papers into it. Then, open more papers, identify common themes, create a corresponding folder, and repeat the process.	Start by reviewing the title, abstract, and author names to organize the categories. Read the first paper to identify the first theme category, upon reading the second, create another category. However, challenges arise in deciding which themes to assign to each paper.	Go through the papers one by one to identify the keywords and design stages for each paper. The keywords or design stages were then used to group the papers.	Identify keywords and design goals, then group papers based on keywords and goals.

4. Results

4.1. Community detection

Figure 1 shows an example network generated from the computer algorithms outlined above. There are a total of 6 communities within that network. A threshold of 0.85 ensures that edges in the network are formed only between nodes with a cosine similarity greater than 0.85. For more information and networks from other years, please refer to the GitHub page (link). Table 3 lists each community label along with the titles of three selected documents from that community.
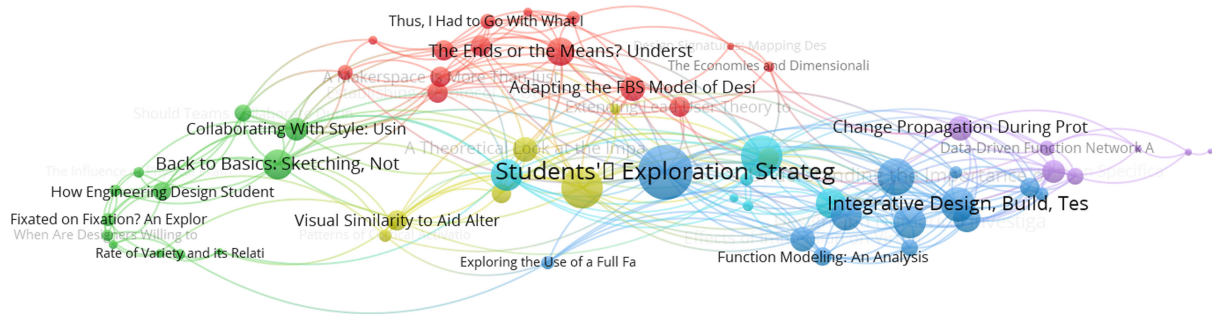

Figure 1. The 2018 network uses Doc2Vec with a threshold of 0.85

According to Table 3, documents from 2018 are categorized into 6 distinct groups. The first group is labeled "Design Methods and Approaches". The titles within the first group encapsulate a variety of innovative strategies and methodologies used in design processes. The label "Prototyping and User-Centered Design" is assigned to the second community which highlights the prototyping strategies in the design process. The third community is labeled "Creative Processes in Engineering Design" because most of the papers in that community focus on creativity and ideation within engineering design contexts. The fourth community is labeled "Influence of Human Behavior on Design" because this community includes papers that discuss how human actions, decisions, and social interactions impact the design process. The fifth community is labeled "Innovative Design Techniques" as this community reflects a focus on novel methodologies and strategies that enhance the design process. The last community is labeled "Function Analysis and Product Design" because papers in this community focus on how understanding functions can enhance design. However, the categorizations are not perfect, and some papers are misclassified. When the papers in a community lack a consistent theme, LLM may assign two topics to a single label to capture the diverse themes within that community. For example, "Prototyping and User-Centered Design" contain both prototyping and user-centered design which are not aligned.

Tabel 3. Community Labels and corresponding three selected document titles

Community labels	Document titles
Design Methods and Approaches	"Design and Numerical Investigation of Rudder With Leading-Edge Protuberances", "Integrative Design, Build, Test Approach for Biomedical Devices With Lattice Structures", "Understanding the Importance of Capturing Climate and Occupancy Trends During Concept-Stage Sustainable Building Design"
Prototyping and User-Centered Design	"Prototyping for Desirability by Design of Experiments: A Case Study of a Hardware Startup", "ContaBat: Designing and Prototyping an Attachable Sports Analytics Device That Provides Ball-Bat Impact Location for Performance Enhancement", "Pulling at the Digital Thread: Exploring the Tolerance Stack Up in Scan to Print Processes"
Creative Processes in Engineering Design	"How Engineering Design Students' Creative Preferences and Cognitive Styles Impact Their Concept Generation and Screening", "Back to Basics: Sketching, Not CAD, Is the Key to Improving Essential Engineering Design Skills", "Fixated on Fixation? An Exploration of the Benefits and Deficits of Design 'Fixation' in Engineering Design"

(Continued)

Tabel 3. Continued.

Community labels	Document titles
Influence of Human Behavior on Design	"The Ends or the Means? Understanding How Students Use Causal and Effectual Information During Design Activities", "Design Sprint for Complex System Architecture Analysis", "Thus, I Had to Go With What I Had: A Multiple Methods Exploration of Novice Designers' Articulation of Prototyping Decisions"
Innovative Design Techniques	"Function Ordering Within Morphological Charts: An Experimental Study", "Inspired Internal Search: Using Neuroimaging to Understand Design Ideation and Concept Generation With Inspirational Stimuli", "Visual Similarity to Aid Alternative-Use Concept Generation for Retired Wind-Turbine Blades"
Function Analysis and Product Design	"Data-Driven Function Network Analysis for Product Platform Planning: A Case Study of Spherical Rolling Robots", "Multi-Level Function Specification and Architecture Analysis Using ESL: A Lock Renovation Pilot Study", "The Function-Human Error Design Method (FHEDM)"

4.2. Human classification of documents

The initial classification of the documents varied in both the number of communities and the corresponding community labels assigned by researchers. To enable a more accurate comparison, the classification results were post-processed, with all communities consolidated into six distinct groups. While the original labels defined by each individual differed, those labels referred to analogous categories, therefore, six standardized labels "design method/processes", "prototyping", "Ideation", "designer's behavior", "design tools", "design goal" were defined to be consistently applied across all communities.

In Figure 2, the cells are colored to indicate communities. The figure illustrates variations in decision-making across individuals. Overall, both congruence and divergence can be observed in the categorizing

Paper	Researcher 1 cluster	Researcher 2 Cluster	Researcher 3 cluster	Researcher 4 Cluster	Researcher 5 Cluster	Computer cluster
Paper 1	design goal	design goal	design tools	design goal	design goal	design method/process
Paper 2	design goal	Designer's behaviour	design goal	design goal	design goal	design method/process
Paper 3	design goal	Designer's behaviour	design goal	design goal	design goal	design method/process
Paper 4	design goal	Designer's behaviour	design goal	design goal	design goal	Designer's behaviour
Paper 5	design method/process	design tools	design goal	design goal	design tools	design method/process
Paper 6	Designer's behaviour	Designer's behaviour	Designer's behaviour	design goal	Designer's behaviour	Ideation
Paper 7	design goal	Ideation	Ideation	design goal	Ideation	design method/process
Paper 8	design goal	design tools	design goal	design method/process	design goal	Prototyping
Paper 9	design method/process	Ideation	design method/process	design method/process	design method/process	design goal
Paper 10	design method/process	Ideation	design method/process	design method/process	design method/process	design goal
Paper 11	Designer's behaviour	design method/process	Designer's behaviour	design method/process	design method/process	design method/process
Paper 12	design method/process	design method/process	design tools	design method/process	design method/process	Designer's behaviour
Paper 13	Designer's behaviour	design method/process	Designer's behaviour	design method/process	design method/process	Designer's behaviour
Paper 14	Designer's behaviour	design tools	design method/process	design method/process	design method/process	Designer's behaviour
Paper 15	Designer's behaviour	design method/process	design method/process	design method/process	design method/process	Designer's behaviour
Paper 16	design method/process	design goal	design method/process	design method/process	design method/process	Prototyping
Paper 17	design method/process	Ideation	design method/process	design method/process	design method/process	Prototyping
Paper 18	Ideation	design tools	design method/process	design method/process	design tools	design method/process
Paper 19	Designer's behaviour	design method/process	design method/process	design method/process	Designer's behaviour	Ideation
Paper 20	design method/process	design method/process	design tools	design method/process	Ideation	design tools
Paper 21	design method/process	Ideation	Designer's behaviour	design method/process	Ideation	Designer's behaviour
Paper 22	Ideation	design tools	design tools	design method/process	Ideation	Designer's behaviour
Paper 23	design goal	Designer's behaviour	design goal	design tools	design goal	design method/process
Paper 24	design tools	design goal	design method/process	design tools	design tools	design goal
Paper 25	design method/process	design tools	design tools	design tools	design tools	design goal
Paper 26	design method/process	design tools	design tools	design tools	design tools	design method/process
Paper 27	Designer's behaviour	design method/process	Designer's behaviour	design tools	Ideation	design method/process
Paper 28	Designer's behaviour	design method/process	prototyping	design tools	Prototyping	design method/process
Paper 29	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour
Paper 30	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour
Paper 31	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour
Paper 32	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Designer's behaviour	Ideation
Paper 33	Ideation	Ideation	Ideation	Ideation	design method/process	design goal
Paper 34	design method/process	Ideation	design tools	Ideation	design method/process	design method/process
Paper 35	Designer's behaviour	Designer's behaviour	Designer's behaviour	Ideation	design method/process	design tools
Paper 36	Designer's behaviour	Designer's behaviour	Designer's behaviour	Ideation	Designer's behaviour	Ideation
Paper 37	Ideation	Ideation	Designer's behaviour	Ideation	Ideation	design goal
Paper 38	design method/process	Ideation	Ideation	Ideation	Ideation	design method/process
Paper 39	design method/process	Ideation	Ideation	Ideation	Ideation	design tools
Paper 40	design method/process	Ideation	Ideation	Ideation	Ideation	design tools
Paper 41	Designer's behaviour	Designer's behaviour	Ideation	Ideation	Ideation	design tools
Paper 42	design method/process	Ideation	Ideation	Ideation	Ideation	design tools
Paper 43	design method/process	Ideation	Designer's behaviour	Ideation	Ideation	Designer's behaviour
Paper 44	design method/process	Ideation	design method/process	Ideation	Ideation	Ideation
Paper 45	design method/process	Ideation	Ideation	Ideation	Ideation	Ideation
Paper 46	design method/process	Ideation	Designer's behaviour	Ideation	Ideation	Ideation
Paper 47	Designer's behaviour	Ideation	Ideation	Ideation	Ideation	Ideation
Paper 48	Designer's behaviour	design method/process	Designer's behaviour	Ideation	Ideation	Ideation
Paper 49	design method/process	Ideation	Ideation	Ideation	Ideation	Ideation
Paper 50	Designer's behaviour	Ideation	Ideation	Ideation	Ideation	Ideation
Paper 51	Designer's behaviour	Ideation	Designer's behaviour	Ideation	Ideation	Ideation
Paper 52	design method/process	Ideation	design method/process	Ideation	Ideation	Ideation
Paper 53	design method/process	prototyping	design tools	prototyping	Ideation	design goal
Paper 54	Designer's behaviour	prototyping	prototyping	prototyping	Ideation	Ideation
Paper 55	design goal	prototyping	prototyping	prototyping	Prototyping	design method/process
Paper 56	design method/process	prototyping	prototyping	prototyping	Prototyping	Designer's behaviour
Paper 57	Designer's behaviour	prototyping	prototyping	prototyping	Prototyping	Designer's behaviour
Paper 58	design method/process	prototyping	prototyping	prototyping	Prototyping	Prototyping

Figure 2. Comparison of document categorization results among five researchers and computer algorithms

decisions. All researchers grouped four papers related to "designer's behavior" together. There is also some agreement on prototyping groupings among researchers 2, 3, and 4. Researcher 1's community configuration differs the most from the other researchers. All researchers show some level of disagreement across all classifications, even in the prototyping grouping, which has the highest level of agreement, the classifications are still not entirely identical.

The differences among humans are already significant, and the gap between computer algorithms and human is even greater. To facilitate a more consistent comparison, the computer-generated topics from 2018 have been further generalized and replaced by human-generated topics. "Design Methods and Approaches" is replaced by "design method/processes". "Prototyping and User-Centered Design" is replaced by "prototyping". "Creative Processes in Engineering Design" is replaced by "ideation". "Influence of Human Behavior on Design" is replaced by "designer's behavior". "Innovative Design Techniques" is replaced by "design tools", and "Function Analysis and Product Design" is replaced by "design goal".

4.3. Topic relationship path

Following the steps outlined in section 3.2, a list of LLM-generated community labels (topics) and their pathways are shown in Figure 3. A threshold of 0.825 indicates that paths are only formed between topics with a similarity score greater than 0.825. The threshold value is determined through iterative testing to better visualize the trends. The behaviors of the topics will be discussed in detail in section 5.
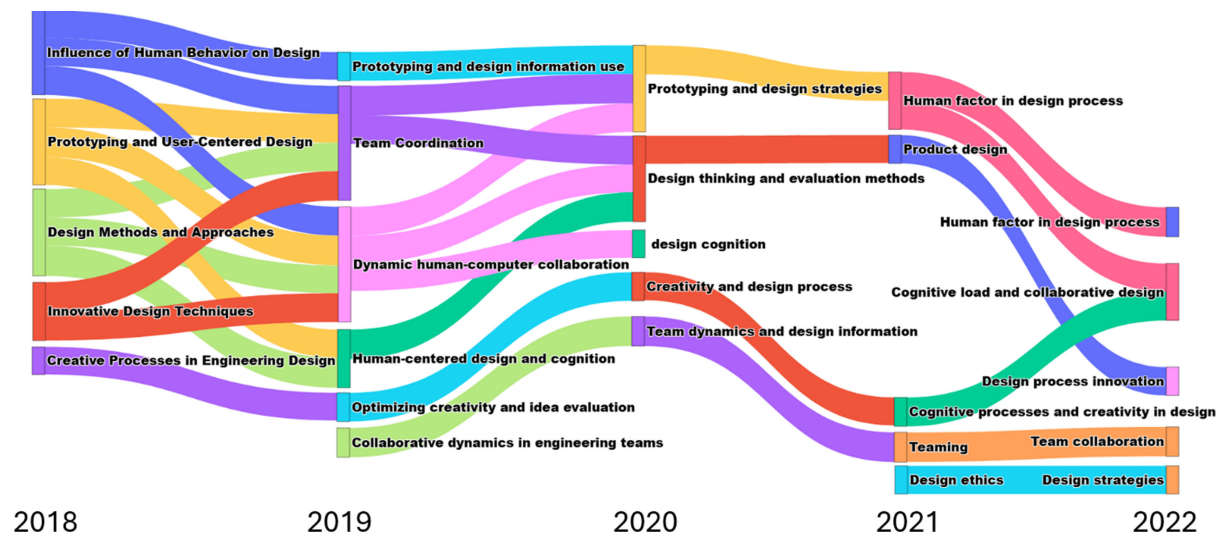

Figure 3. Sankey diagram from Doc2Vec model with a threshold of 0.825

5. Discussion

5.1. Comparison of communities identified by human and algorithm

The comparison between human and computer algorithms demonstrated a significant degree of variability in classification results. It is observed that only "ideation" and "designer's behavior" show some level of agreement between computer-generated output and human-determined output. Such an agreement is predictable since papers that discuss ideation and designer's behavior are quite differentiable than papers under other themes due to the unique terminologies they inherently involve. On the other hand, the differences may be due to the different focuses during the decision-making process. For example, while computers categorize documents entirely based on document similarity, which is a summation of word similarities within each document, the human classification process often involves intuition and is based on divergent reasoning. Therefore, documents containing identical words or words with similar meanings are likely to be grouped by computers, even if these papers focus on different aspects.

One limitation of the study is that parameters still have impacts on the results. A small change in the threshold can change the network structure significantly. A key reason for the non-robust behavior is that

the cosine similarity across most vector pairs shows little variation. In such a case, the community detection algorithm does not work well. Future research that involves integrating more advanced machine learning such as graph-neutral networks to improve performance, and incorporating manual labor where algorithms improve with human feedback should be explored.

5.2. Research trends

There is some merging, splitting, and discontinuation observed in Figure 3. For example, the topics "Influence of human behavior on design", "Prototyping and user-centered design", "Innovative design techniques," and "Design methods and approaches" from 2018 merge into a single topic, "Dynamic human-computer collaboration," in 2019. This merging suggests that the research focus has shifted from understanding human behavior, developing new design methods, and studying prototyping strategies into a combined area that focuses on creating dynamic collaborations between human and computer. Additionally, these same topics from 2018 also influenced other areas in 2019. For instance, the same four topics also flow into the 2019 topic of "Team coordination". These additional flows indicate that some topics from 2018 not only combine into new areas but also branch out, impacting the development of multiple research directions. Several discontinuations are also observed, which means that topics from one year do not connect to any topics in the subsequent year, such as "design cognition". The discontinuity suggests that certain areas of research temporarily lost prominence.

There are also a few paths without merging or splitting behaviors, where a single, direct flow connects one topic to the next cross years. In these cases, the topics remain consistent over time. For instance, a path that starts with "Creative processes in engineering design" and ends with "Cognitive load and collaborative design," and the path that starts with "Collaborative dynamics in engineering teams" and ends with "Team collaboration" maintain a steady focus on creativity and teaming, respectively, which suggests that creativity and teaming are stable, well-established areas with ongoing research interest and relatively little change in focus.

6. Conclusion

This study developed a framework for visualizing the research trends within the design methodology community. By applying additional models to the same dataset and employing LLM to determine topics, this study extends our previous research to gain deeper insights into the topic relationships. Although the Sankey diagram demonstrates limited effectiveness in visualizing the trend, it highlights several phenomena, such as the strong correlation between most topics, with only a few topics being relatively independent and minimally influenced by other topics. A limitation of our study arises from the poor performance of the document categorization method. Although our study involves human validation by comparing how humans and computer algorithms categorize documents, the difference is substantial. The disagreement between human and machine-generated results indicates the need for further investigation into the factors influencing community categorization and incorporating more advanced techniques to improve the results.

References

Chen, H., Wang, X., Pan, S., & Xiong, F. (2021). Identify Topic Relations in Scientific Literature Using Topic Modeling. *IEEE Transactions on Engineering Management*, 68 (5), 1232–1244. https://doi.org/10.1109/TEM.2019.2903115

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Hagberg, A.A., Schult, D.A., Swart, P., & Hagberg, J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the Python in Science Conference*, (pp. 11–16). http://dx.doi.org/10.25080/TCWV9851

Huang, L., Chen, X., Zhang, Y., Wang, C., Cao, X., & Liu, J. (2022). Identification of topic evolution: network analytics with piecewise linear representation and word embedding. *Scientometrics*, 127 (9), 5353–5383. https://doi.org/10.1007/s11192-022-04273-1

Le, Q., & Mikolov, T. (2014). Distributed Representations of Sentences and Documents. In P. X. Eric & J. Tony (Eds.), *Proceedings of the 31st International Conference on Machine Learning*, 32 (2) (pp. 1188–1196). PMLR. https://proceedings.mlr.press/v32/le14.html

Leydesdorff, L., & Nerghes, A. (2017). Co-word maps and topic modeling: A comparison using small and medium-sized corpora (N < 1,000). *Journal of the Association for Information Science and Technology*, 68 (4), 1024–1035. https://doi.org/10.1002/asi.23740

Lo, S.-C., & Wang, C.-C. (2023). Revisit Girvan-Newman Algorithm for Research Topic Analysis: An Application on Library and Information Science Studies. *Journal of Library & Information Studies*, 21 (1), 1–16. https://doi.org/10.6182/jlis.202306_21(1).001

Mehta, V., Bawa, S., & Singh, J. (2021). WEClustering: word embeddings based text clustering technique for large datasets. *Complex & Intelligent Systems*, 7, 1–14. https://doi.org/10.1007/s40747-021-00512-9

Mu, W., Lim, K. H., Liu, J., Karunasekera, S., Falzon, L., & Harwood, A. (2022). A clustering-based topic model using word networks and word embeddings. *Journal of Big Data*, 9 (1), 38. https://doi.org/10.1186/s40537-022-00585-4

van Eck, N. J., & Waltman, L. (2014). Visualizing Bibliometric Networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact: Methods and Practice*, (pp. 285–320). Springer International Publishing. https://doi.org/10.1007/978-3-319-10377-8_13

Vayansky, I., & Kumar, S. A. P. (2020). A review of topic modeling methods. *Information Systems*, 94, 101582. https://doi.org/10.1016/j.is.2020.101582

Wang, P., Xu, B., Xu, J., Tian, G., Liu, C.-L., & Hao, H. (2016). Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification. *Neurocomputing*, 174, 806–814. https://doi.org/10.1016/j.neucom.2015.09.096

Xiao, S., McAdams, D. A. (in press). Employing network analysis to identify research topic evolution, *Proceedings of the Design Society*, 5.

Zhang, T., Cui, W., Liu, X., Jiang, L., & Li, J. (2023). Research on Topic Evolution Path Recognition Based on LDA2vec Symmetry Model. *Symmetry*, 15 (4), 820. https://www.mdpi.com/2073-8994/15/4/820